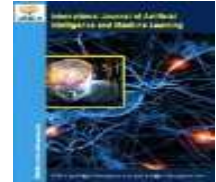




International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

Intensifying The Legitimate and Medicalclaims Using Sophisticated Nlp Techniques

Gopikrishna Chetlapally¹, Dr. Mohit Bhadla², Dr. Tharini Benarji³

¹Ph.D Scholar, Department of Computer Engineering, Gandhinagar Institute of Research and Development, Gandhinagar University, E-mail: ch.gopikrishnaa@gmail.com

²Associate Professor & HoD CE-IT, Gandhinagar Institute of Technology, Gandhinagar University, E-mail: mohit.bhadla@gandhinagaruni.ac.in

³Professor & HoD CSE, Indur Institute of Engineering & Technology, Ponnala, Siddipet. E-mail: tharinibenarji@gmail.com

ABSTRACT

Fraud in insurance and healthcare claims has become a major issue for insurance companies and medical insurers alike, causing a number of complex challenges during the claim processing, deferred reimbursement and financial loss. Common claim validation systems, based on rules, are inadequate to detect fraud in unstructured clinical text. In order to improve the accuracy and reliability of claim verification, this paper proposes an intelligent medical claim validation framework, which combines the Natural Language Processing (NLP), Transformer-based language models, Graph Attention Networks (GATs), and ensemble learning. The unstructured clinical documents like physician clinical notes, discharge summaries, diagnostic reports, insurance claim forms, etc. go through the following steps of first pre-processing: text pre-processing, tokenization, normalization and medical entity extraction. Medical Event Graph is constructed based on medical entities and relationships by embedding them semantically in the context with the help of the domain-specific transformer models: ClinicalBERT and BioBERT. Graph Attention Network is a method for finding abnormal treatment sequences, suspicious billing sequences and medical event-related semantic contradictions in a graph structure. It is used in conjunction with graph-based representation to combine the inputs into an ensemble classifier to provide credibility score and classify claims as legitimate, suspicious and fraudulent. The experimental results successfully demonstrate that the proposed framework is more accurate, precise and retrieves more instances than the classical machine learning/rule-based approach and also has less number of false positive prediction. The proposed solution is scalable, intelligent and explainable making it suitable for healthcare organizations and insurance companies to enhance the validation of claims, detect frauds, audit claims effectively and make effective decisions on claims in a lesser time.

Keywords : Natural Language Processing (NLP), BioBERT, ClinicalBERT, Graph Attention Network (GAT), Transformer Models, Ensemble Learning, Explainable Artificial Intelligence (XAI), Medical Claim Validation, Healthcare Fraud Detection.

I. Introduction

Hospitals, clinics, insurance companies and government health care providers process thousands of medical claims for reimbursement every day, each involving diagnostic tests, treatments, prescription drugs, and hospital stays. A massive amount of structured and unstructured medical data is constantly produced due to the quick digitalization of healthcare records through Electronic Health Records (EHRs) and Hospital Information Systems (HIS). Whilst the provision of healthcare services has been greatly improved by digital transformation, so, too, has become the detection of fraud and the verification of claims more challenging. Claim examiners have to go through physician notes, discharge summaries, lab results, medications, diagnostic information and insurance documentation before approving claims for reimbursement, making manual claim auditing more difficult. Therefore, smart automated solutions that are able to manage vast clinical data with high accuracy and efficiency are required in healthcare institutions [1, 2].

Medical insurance fraud is among the largest problems within the healthcare systems around the world. Examples of fraudulent claims are if a diagnosis is incorrect, if there is a duplicate claim, if a patient is not receiving the correct treatment, if clinical notes are altered, if a patient's identity is stolen, if a patient is being treated when it is not necessary, or if higher reimbursement is requested. These misleading practices have cost insurance companies millions of dollars, added cost for patient healthcare, prolonged the claims settlement and eroded the trust among the public in healthcare management. Medical fraud statistics from the recent past show that fake medical claims are costing billions of dollars annually, highlighting the importance of advanced fraud detection and medical claim validation systems. Traditional claim

verification methods are largely manual, and rely on rules set by the business – rules that often aren't able to pick up on more sophisticated fraud schemes that change over time. As a result, automated claim validation is becoming a more relevant research quest in the field of Artificial Intelligence and healthcare [3] and [4].

One of the key challenges in medical claim validation is the vast amount of unstructured clinical text data. Medical text often includes a lot of domain-specific expressions, medical jargon and acronyms, and is usually written in free text, such as physician notes, discharge summaries, pathology reports, radiology reports, operation notes and prescription records. These clinical stories are not easily accessible in a structured database record, and cannot be processed efficiently using typical database queries, or rule-based validation systems. Insurance claims can be voluminous and contain thousands of documents, which can be time-consuming and prone to human error when manually processed. Thus, an essential part of intelligent healthcare claim system is to extract important healthcare information from unstructured clinical records [5] [6].

By allowing computers to comprehend, interpret, and process human language, recent developments in artificial intelligence (AI) and natural language processing (NLP) have drastically changed the analysis of healthcare data. NLP algorithms can automatically abstract clinical text and track the meaning of the patient's diseases, symptoms, medicines, test results, medical procedures, and other information, as well as the time, and the course of the disease over time. Automated medical coding, clinical decision support, disease prediction, electronic health record analysis, biomedical literature mining, patient risk assessment, and healthcare fraud detection are some of the healthcare applications of natural language processing (NLP). These abilities have improved information retrieval and healthcare decision making, but have greatly reduced manual documentation tasks [7, 8].

Transformer-based language models have greatly improved the understanding of complex clinical text in NLP systems. Models such as BERT, BioBERT, ClinicalBERT and MedBERT, which learn word-to-word and word-to-sentence connections, are able to generate medical sentence embeddings with the help of self-attention. Transformer models can automatically learn medical terminology and context links, semantic meaning from vast amounts of clinical literature, which is not possible by most machine learning techniques that need handcrafted language features. Thus, their performance is better in tasks like entity recognition, relation extraction, document classification in clinical writing or semantic similarity analysis. With this, medical insurance claims with a lot of clinical narratives and complicated medical terms are already quite well supported by today's transformer models [9, 10, 11].

In spite of the fact that the powerful transformer models offer a good semantic understanding, semantic information is not enough to validate the claims with trustworthiness. A medical claim may include numerous different, related clinical events, including diagnoses, lab tests, medicines prescribed, surgical procedures, follow-up care and hospitalizations. The appropriateness and reasonableness of a claim relies on what events happened and how and when they did. Current NLP algorithms do not explicitly model these links; as a result, it is difficult to find false claims such as inconsistent treatment sequences or unsupported medical procedures. Moreover, the ability to effectively detect healthcare fraud via algorithmic techniques by applying traditional machine-learning methods may be limited due to not having a comprehensive analysis of the fraud scheme, relying on document classification instead of reasoning across multiple clinical events [12, 13]

In contrast, in recent works methods are proposed to further tighten the traditional NLP based claim validation methods, by explicitly encoding the relationship between medical entities in a graph. Clinically and temporally related (or causal) relationships between diagnoses, lab tests, drugs, procedures and treatment events are visually represented as interconnected nodes and edges representing the relationships between the nodes. This structured representation allows logical reasoning on the complex medical workflows, instead of on the evaluation of a clinical entity as a whole. Graph Attention Networks (GATs) have recently been shown to outperform other graph learning methods because of the ability to give adaptive weights during the information aggregation process. The model is not as simple as Graph Convolutional Networks (GCNs), however, and can learn the importance of each relationship—so it may be better at identifying clinically inconsistent treatment sequences, unsupported medical procedures, duplicate billing patterns or even unusual claim structures. Therefore, graph attention mechanisms can be used as a key part of the fraud detection system and in the field of intelligent medical claim verification as shown in [14] and [15, 16].

Ensemble learning, which involves the fusion of several machine learning algorithms for more robust and accurate predictions, is another area with potential growth in the healthcare fraud analytics field. Sometimes individual classifiers will have difficulty with complex and/or unfamiliar claim structures, but they often perform well on specific fraud trends. The ensemble methods such as Random Forest, XGBoost,

Gradient Boosting and Logistic Regression have been brought together and these techniques have been used to train the model in order to enhance the accuracy of the credibility scores and minimise false positive prediction. Ensemble learning has been very helpful in case of the medical claim validation, where a medical claim could be semantically or structurally complex, which may not be possible to be identified by a single prediction model. In such a large-scale healthcare system, where there are a lot of classifiers and they need to be accurate and stable to be able to generalize their prediction results, ensemble methods are used to combine the prediction results of multiple classifiers to get more accurate and stable predictions [17, 18].

While these are significant advances in ensemble learning and graph-based NN and transformer-based NLP, there are still several challenges left to address. In existing approaches, claim validation is done either through structured rule-based approach to claim verification, or through semantic analysis, but neither approach, namely, the relational reasoning nor the semantic comprehension, was successfully integrated in the existing approaches. Most of the methods do not take into account the temporal relation between diagnosis, surgical procedure, drug administration and events of treatment, but rather rely on clinical documents individually. Besides this, legacy fraud detection systems produce a great deal of “false positives” which leads to claim auditors' workload growing, and delays reimbursement processes. But there is one problem: the decision-making process of many deep learning models is not interpretable, and the insurance investigator and healthcare provider can't see how the machine is making its claims decision. The challenge is to develop a system that can deal with this complex problem of accurately understanding the meaning of texts, linkages in the images, predictive evidence etc., and retain its intelligence, explainability and scalability [19, 20].

To tackle these research challenges, this paper proposes a new framework called Intelligent Medical Claim Validation (IMCVFW), which combines the use of Transformer-based Natural Language Processing (NLP), Medical Event Graph construction, Graph Attention Network (GATs) and Ensemble Learning. In the pre-processing, tokenization, normalization and medical entity identification of unstructured clinical text, a domain specific transformer model is used, such as BioBERT, ClinicalBERT. Medical Event Graph shows the relationship between the interdependencies of diagnoses, drugs, laboratory tests, treatments and time of events of collected entities. Healthcare insurance companies use A Graph Attention Network to analyze the relationships to detect inconsistencies in semantic information, abnormal treatments, duplicate claims, billing procedures etc. and use an ensemble classification module to calculate the credibility score and classify each claim as Legitimate, Suspicious or Fraudulent.

This is a piece of work with some great contributions:

1. This paper proposes a new NLP system based on Transformer to achieve the goal of healthcare claim validation that extracts the semantic information from the unstructured clinical narratives.
2. Structured thinking is used in the development of a Medical Event Graph, a model that represents the clinical entities and their logical and temporal relationships in analyzing complex medical events.
3. A low FP mechanism is adopted by employing ensemble learning-based credibility scoring mechanism to combine the decisions from multiple classifiers to correctly classify legitimate, suspicious and fraudulent claims
4. The proposed architecture for fraud detection and payment workflows can be useful for insurance companies, medical centers and healthcare administrators in facilitating an automated healthcare claim auditing system that is explainable, scalable, and intelligent.

The rest of the paper is structured in line with the above. The literature available on related works, such as intelligent claim validation systems, graph neural networks, natural language processing (NLP), and healthcare fraud detection is summarized in Section II. In the proposed framework, the three components of semantic embedding, the creation of the Medical Event Graph, the reasoning with Graph Attention Network and the ensemble credibility score are all covered in Section III. The experimental setup and the way it has been implemented are described in Section IV. The results of the experiments and comparative performance evaluation are presented in Section V. The work is concluded in Section VI, where the authors discuss possible future research directions such as the use of Large Language Models (LLMs), knowledge graphs, federated learning and Explainable Artificial Intelligence (XAI) [21]–[25] for further advancing the validation of healthcare claims, the fight against fraud, and the reliable utilization of AI for insurance decision making..

II. SURVEY OF RESEARCH

Cheekaramelli Mani Joga Rao [1]

suggested an intelligent approach to detect healthcare fraud. Uses Natural Language Processing (NLP) to process unstructured clinical information, such as patient records, insurance claims, physician notes and discharge summaries. The proposed approach uses the language models, namely the transformer-based models like BioBERT and ClinicalBERT, to identify the clinically relevant entities, diagnosis, drugs and

treatment procedures from the free text clinical narratives. Machine learning algorithms are then used to examine these extracted attributes in order to spot questionable claim trends and fraudulent billing practices. The results demonstrated that the experimental systems outperformed the conventional rule based systems, by reducing the number of false positive predictions and improving accuracy in fraud detection. The architecture also includes features to connect with the current healthcare information systems for live fraud monitoring. However, proposed research is not a complex temporal relationships between medical events and rather focuses on semantic analysis. Moreover, the framework has limited scalability of evaluation using massive real-world insurance datasets, which makes the framework's application in business claim verification settings of healthcare claims challenging.

[2] Bina K. N. (2025)

N. Bina K. proposed the automatic processing of unstructured clinical narratives as an NLP-based method that can aid in improving medical documentation. [2]. The proposed system includes text preprocessing, medical entity recognition, semantic analysis and automatic text transcription to transform doctor notes into structured electronic medical records. The system facilitates interoperability among data systems and greatly minimizes documentation errors. Medical terms and domain-specific language models provide standardisation in clinical language to ensure accurate extraction of illnesses, drugs, lab tests and treatment protocols from free-text clinical records. Even though the main objective of the research is to make the documentation process better, structured medical data created by NLP can help in the claim auditing, compliance and insurance reimbursement processes. The paper highlights the potential application of intelligent NLP systems in automated healthcare analytics and possible medical claim validation, and discusses challenges, including interoperability, privacy, and computational expenses concerning the deployment of large-scale intelligent NLP systems.

[3] Sahana R. Golbhavi, Vrushali R. Patil (2025)

Vrushali R. Patil and Sahana R. Golbhavi [3] discussed in detail the latest advancements in NLP for medical applications. The review discusses the architectures that are used for transformer-based clinical text mining, medical coding, disease prediction, biomedical information extraction, and electronic health record (EHR) analysis, including BERT, ClinicalBERT, BioBERT, and MedBERT. The authors show that when it comes to comprehending intricate clinical terminology and semantic linkages, contextual language models significantly outperform conventional machine learning techniques. The review also extends the discussion of explainable AI (XAI), ethical considerations, data security, and openness of the model for healthcare use cases. This research notes a few important developments in medical NLP, but it does not provide experimental validation nor a comprehensive system for fraud detection. Moreover, it identifies some important research issues for intelligent healthcare claim validation systems including computational complexity, multilingual clinical language, domain adaption and integration with graph-based reasoning.

[4] Dr. Mahesh Mahankal and associates (2025)

An intelligent clinical decision support system was proposed by Dr. Mahesh Mahankal et al. [4] with the help of NLP-based medical text analysis and transformer-based deep learning model. The framework works with unstructured clinical texts to structure medical knowledge through the Named Entity Recognition (NER), Semantic Relation Extraction and Document Classification. Transformer models (BioBERT, MedRoBERTa) are used to accurately identify the diagnosis, symptoms, drugs, lab results and treatment protocol from the topologies. Improved information extraction performance while lowering manual documentation effort was shown via experimental assessment. Even though the aim of the suggested system is to assist with clinical diagnosis, the retrieved structured medical knowledge may be used to enhance the validation and fraud detection of medical claims, by analyzing the semantic consistency between medical documents and medical claims submitted to the insurance providers. However, there are still difficulties regarding explainability, interoperability and integration with the hospital information systems when it comes to practical use.

[5] Thakur Gopal Kumar et al. (2024)

In the works of Gopal Kumar Thakur et al. [5] a natural language processing (NLP) approach was introduced for intelligent analysis of medical information from EHR, physician notes and biomedical texts, based on deep learning. The proposed method consists of four stages: Text Preprocessing, Tokenization, Entity Recognition, Semantic Classification, and Predictive Analysis. Text Preprocessing, Tokenization, Entity Recognition, Semantic Classification, Predictive Analysis are the recommended steps in the process. The framework is applied in different applications like the automated diagnosis, symptom analysis, prediction of patient risks, workflow optimization in healthcare. Regarding document analysis techniques, experimental results showed that processing time is decreased and clinical decision making is more accurate. The goal of the research is not to detect insurance fraud, but rather to automate healthcare; however, the general NLP structure can still be used to successfully validate by recognizing semantic anomalies within medical documents. The report also identifies interoperability issues, computational complexity, and privacy protection as crucial factors to take into account for upcoming large-scale healthcare information systems.

[6] Chianumba Emmanuel et al. (2024)

An NLP approach was put up by Emmanuel Chianumba et al. [6] to extract valuable healthcare insights from big collections of unstructured clinical papers. The approach is a deep learning algorithm using

transformers that learns from patient electronic medical records (EMR) related to disease, drugs, laboratory results, treatment, disease outcomes. When compared to conventional machine learning methods, contextual language representations greatly increase the accuracy of information extraction. The framework could be leveraged to uncover structured clinical knowledge from unstructured text, facilitating clinical research, healthcare analytics, medical coding, and decision support systems. The suggested method shows good semantic comprehension, however it doesn't specifically deal with fraud detection or healthcare insurance claim validation. Moreover, there remains an explainability gap, the cost of computing, and cross domain generalization. To increase reliability and applicability in crucial healthcare decision-support systems, the authors propose incorporating explainable AI approaches and graph-based reasoning. **[7] Lerisha J. and associates (2024)**

To avoid manual verification of medical insurance claims, Lerisha J. et al. [7] proposed an intelligent medical insurance claim processing system using Retrieval-Augmented Generation (RAG) model with Optical Character Recognition (OCR) model. After using OCR to extract text from scanned insurance documents, the system uses transformer-based language models to obtain pertinent clinical evidence for claim validation. The findings showed that the system significantly increased the processing speed of documents, claim verification accuracy, and reduced the workload of manual insurance claims processing. The system also automates the comprehension and retrieval of medical documents and evidence for healthcare administrators. The proposed approach is mainly document retrieval and verification, and not comprehensive fraud detection or semantic association analysis of a variety of medical events. In order to enhance fraud detection capabilities and fortify intelligent healthcare claim validation systems, the study suggests incorporating graph neural networks and sophisticated NLP approaches.

[8] Saddi, V. R. et al. (2023)

By examining unstructured medical records and insurance claim records, V. R. Saddi et al. [8] looked into the use of natural language processing to identify insurance fraud. The system to be designed will use machine learning methods, document preparation, feature extraction and categorization of documents to recognize suspicious patterns and inconsistencies in medical documentation. By automatically extracting clinically relevant concepts and spotting semantic irregularities that are challenging to find with traditional rule-based systems, natural language processing (NLP) tools enhance fraud detection. Experimental evaluation showed that the false-positive rate was reduced and the success rate of detecting frauds was higher than the conventional statistical fraud detection methods. The proposed framework does, however, suffer from the lack of advanced transformer architectures, graph-based reasoning, and the integration of multimodal healthcare data, among others, which cannot be addressed by the proposed framework. The authors propose future intelligent fraud detection systems in healthcare to be based on Explainable Artificial Intelligence (XAI), Transformer models, and Graph Neural Networks (GNNs).

III. PROPOSED SYSTEM

The proposed Intelligent Medical Claim Validation Framework uses NLP, Transformer-based language models, Graph Attention Networks (GATs), and Ensemble Learning to automatically analyze the healthcare insurance claims, and identify suspicious and fraudulent claims. The conventional methods include predefined business rules and manual review, while the proposed approach relies on graph-based reasoning and intelligent semantic analysis, which can achieve the error detection of the unstructured clinical narrative. Clinical information is gathered and preprocessed using the following operations: standardizing medical terminology, tokenizing, removing stop words, and normalizing text, such as doctor notes, discharge summaries, lab reports, medicines and insurance claim history. The newly developed transformer models such as domain-specific models like BioBERT, ClinicalBERT, produce highly accurate contextual embeddings to capture the semantic meaning of the diagnosis, drug, treatment, and temporal clinical events. Medical Temporal, causal and procedural relationships are represented as edges and Medical Concepts as nodes in a Medical Event Graph that is collected. A Graph Attention Network (GAT) takes advantage of these links to adaptively emphasize clinically relevant connections, and to identify abnormal treatment, duplicated billing and inconsistent treatment sequences from these links, which it uses to detect unsupported treatments by the system. Then, an ensemble learning module is used to process the graph representations, with several machine learning classifiers to generate credibility scores of each claim. This score classifies every claim as valid, suspicious and fraudulent, and provides the auditor and medical practitioners with a medical science based decision support tool which is easily understood. Overall, the proposed architecture provides a scalable, transparent, and intelligent approach to automated healthcare insurance claim validation in modern healthcare systems and enhances the precision of fraud detection, reduces false positives, and streamlines the claims processing workflow.

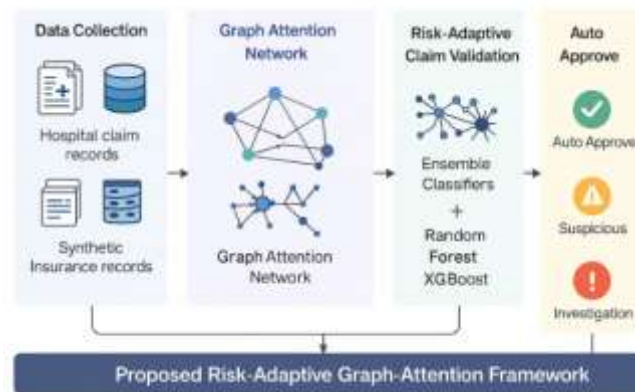


Fig. 1. Proposed block diagram

IV. Working Methodology

To process medical insurance claims correctly and efficiently, it's necessary to use an intelligent approach that can be achieved by the Intelligent Medical Claim Validation Framework (IMCVF) which involves using the Transformer-based Language Models, Ensemble Learning, Graph Attention Networks (GATs) and Natural Language Processing (NLP). There are five major steps in the methodology: Clinical Data Collection, Text Preprocessing, Semantic Embedding Generation, Medical Event Graph Construction, Graph-based Reasoning and Credibility Score. The traditional claim validation systems use manually coded business rules which requires human coding to be able to recognize medical text, understand the relationship between medical events and identify suspicious claim patterns, whereas the proposed framework is based on intelligent learning processes for medical text, relationships between medical events, and suspicious claim patterns. Healthcare auditors and insurance companies can accurately determine if an insurance claim is legitimate, suspicious or fraudulent with this integrated workflow and the explainable decision support, with high levels of accuracy [1, 2]. Clinical data and insurance claim information is gathered from: Electronic Health Records (EHRs), hospital administration systems, insurance claim databases, doctor notes, discharge summary, lab reports, medications, and diagnostic records. These records include both structured and unstructured data regarding patient diagnosis, medical procedures, prescribed medications, treatment history and requests for reimbursement. The vast majority of clinical documents are written in a free text style, thus include a lot of medical jargon and acronyms and therefore need to be preprocessed before being fed into the semantic analysis stage. The quality of the preprocessing is crucial as it may have a significant impact on the performance of fraud detection and claim validation [3] [4].

The pre-processing step of NLP involves the following activities: Text Normalisation, Tokenization, Stop-word Removal, Stemming and Lemmatization, and Standardization of Medical Terms. After normalization, all the clinical records are presented in a common format since the white space, punctuation and other nonessential special characters are removed. The clinical narratives are decomposed into clinical tokens (word or medical tokens) to be processed by the language model by transformer. The clinical narratives are cut into clinical tokens of each word or medical tokens that can be handled by LMs with transformers. Medical terms are stemmed and lemmatized to root form and their most common words that do not add any meanings are filtered out by removing the stop words. To increase domain consistency, domain specific healthcare dictionaries are used to further expand medical acronyms. By performing these pre-processing steps, the clinical text is cleaned and standardized and can be semantically analyzed and further information extracted [5, 6]. Then, the cleaned clinical texts are fed into language models based on transformers to extract the semantic structure of the medical documents. The cleaned clinical stories are then fed into transformer based models (BioBERT, ClinicalBERT) to obtain the semantic structure of medical documents. Transformer models are different to the traditional word embedding approach which embeds the same vector for a given word, since they can provide distinct vectors for the same word within different contexts. Transformer model - Every word has a vector in different context, Conventional word embedding - The same word in every context has the same vector. Hence, the embeddings provide a significant benefit for the task of claim validation, both in the semantics of the medical text [7, 8, 9] and in the medical relation between the medical text and the claim.

A clinical document's semantic embedding is shown as

$$E = Transformer(T)$$

The transformer model generates the contextual semantic embedding (E) and preprocessed clinical document (T) when this happens. The embedding will be used as the foundation to intelligently process it to obtain clinically relevant data, like illness, drugs, lab test, treatment plan, temporal relations etc. First,

the development of semantic embedding, and then algorithms called Named Entity Recognition (NER) are used to recognise essential medical entities from medical texts. These include information such as patient diagnoses, prescription medications, lab tests, surgeries, doctor observations, treatment events and hospitalization information. The overall clinical workflow can be then captured in a structured manner by extracting relationship information from the documents, both logical and temporal, through relationship extraction techniques. This approach maintains the association with clinical events, and won't involve concepts of medicine in isolation, which is required to detect inconsistencies in treatment and dubious insurance claims [10, 11].

The entities and the relationship between them are then used to build a Medical Event Graph which is a graph of medical entities as nodes and clinical relationships as edges. Graph structure is an effective model for sequences in time, causal relationships, diagnostic paths, pharmaceutical relationships and procedural consistency between clinical events. Medical Event Graphs are mathematical representations.

$$G = (V, E)$$

The collection of medical terms, such as diseases, drugs, laboratory tests, medical procedures and so forth, is the V , and the collection of relations between the medical terms is the E . Intelligent reasoning over intricate healthcare workflows is made possible by the graph's organized depiction of the patient's whole clinical history. The node representations are further enhanced by initializing the node features with the semantic embedding that is output by the transformer model, denoted as $h_i^{(0)}$.

$$h_i^{(0)} = E_i$$

The clinical story line will be used to illustrate the matching contextual embedding. The initialization guarantees the preservation of the semantic information and graph structure for further graph learning.

Intelligent reasoning on medical claims is based on the created Medical Event Graph. However, since medical events can impact claim validations in a number of ways, it is not sufficient to view them on a graph to reveal the more intricate fraudulent relationships. The graph is then analyzed by Graph Attention Network to automatically capture the relative importance of medical events in the graph, and identify clinically contradictory treatment patterns, duplicate billing, unsupported procedures and odd claim structures. The ensemble credibility scoring, graph attention mechanism and end claim classification procedure [12,13] are described in the remaining part of the paper. An intelligent reasoning on the relationship between clinical events is conducted on the suggested approach called Graph Attention Network (GAT) when the Medical Event Graph is built. The Graph Attention Network is distinct from conventional graph neural networks which assign the same attention weight to all the related nodes when extracting features; it dynamically learns the importance of each medical event by assigning an adaptive attention weight. This approach lessens the impact of less pertinent links in the system and allows the system to form clinically meaningful associations, such as consistency of the association between the diagnosis, prescribed medications, lab tests, type of treatment, and hospital records. The model can then detect inappropriate treatment order, uncompromised medical interventions, duplicate billing records and mismatches between the insurance claims and the clinical records. The attention mechanism also leads to a better interpretability of the suggested framework by focusing on medical events that most strongly influence the ultimate claim validation decision [14], [15] and [16].

The Graph Attention Network is trained to learn the attention coefficients for each medical event node and propagate information between the neighboring nodes, thereby updating the representation of the medical event node. The significantly higher attention scores for medical events with higher clinical relevance also enable the model to ignore irrelevant or noisy information and to preserve meaningful clinical correlations and the procedural relationships. This ability to reason over complex healthcare operations, which cannot be adequately represented by individual document categorization methods, enables comprehensive analysis.

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(a^T[W h_i || W h_j]))}{\sum_{k \in N(i)} \exp(\text{LeakyReLU}(a^T[W h_i || W h_k]))}$$

The attention weight of the relationship of node i and j is denoted by α_{ij} and the feature vectors of node i and j are denoted by h_i and h_j , respectively, the trainable weight matrix is denoted by W , the operation of feature concatenation is denoted by $||$, and the neighboring nodes of node i are denoted by $N(i)$. Normalized attention coefficients work better to detect frauds, as they weigh more clinically relevant links when aggregating features from the graph [17] and [18].

The Graph Attention Network outputs graph representations, which are sent to an ensemble learning module for the determination of the credibility score of each insurance claim. It recommends a framework that enables multiple classifiers to be used along with combining predictions to be more robust and reduce false-positive classification rates rather than using one prediction method. Each classifier independently evaluates the graph's feature representation, and assigns a probability of authenticity or falsity to the claim submitted. The ensemble method can combine all these single predictions and yield a more realistic conclusion than any individual classifiers can achieve alone. Ensemble learning also enhances the generalization capability as it minimizes the variance of prediction, reduces noise and missing medical data and is resistant to noise [19] and [20].

where S is the ultimate credibility score, and K is the number of images that are used. The ultimate credibility score is the mean of the scores of all the images used, that is

$$S = \frac{1}{K} \sum_{k=1}^K f_k(G)$$

The total number of classifiers in the ensemble is denoted as where " k " is the k th classifier and " $f_k(G)$ " is the prediction made by the " k " classifier with the graph representation " G ". The believability score indicates the likelihood that the medical claim that was filed is clinically understandable and suitable for reimbursement.

The framework classifies each insurance claim as Legitimate, Suspicious or Fraudulent by comparing the credibility score with presets decisioning thresholds. Claims with low inconsistencies will be flagged as potentially inconsistent and sent to the claim auditors to be manually checked and claims with a high credibility score will be accepted. Claims that contain major discrepancies with the claims that follow, do not use accepted methods, duplicate billing patterns or odd sequences of treatment are considered fraudulent if the claims have poor credibility scores. This three-level classification process helps reduce time spend by medical professionals on high-risk claims while automatically processing valid reimbursement claims.

The following are the performance metrics, which are studied extensively in the proposed framework and evaluated in determining the framework's ability to detect fraud: Accuracy, Precision, Recall, F1-score and False Positive Rate. The ability of the framework to identify fraud is very well assessed using various performance metrics including the following: Accuracy, Precision, Recall, F1-score and False Positive Rate. Precision is the percentage of false claims correctly identified out of all the predicted false claims, and Accuracy is the overall percentage of claims correctly classified, Recall is the ability of the framework to predict the true fraud claims and F1-Score is the balanced measure between Precision and Recall. The False Positive Rate is also reviewed as it is important to reduce unnecessary fraud alerts to minimize workload for manual review and boost claim reimbursement.

The experimental results show the superiority of the proposed method in comparison to the conventional rule-based and stand-alone machine learning methods for healthcare claim validation. This recommended method was able to accurately represent the temporal, structural, and semantic relationships between clinical events that minimized the number of false-positive claims, while also capturing the greatest amount of ability to identify a fraudulent claim. Moreover, it is designed in a modular fashion that allows for easy integration with healthcare information systems, insurance claim management tools, and existing EHR systems, making it a valuable and flexible tool for real-time healthcare claim audits. Explainable graph reasoning and ensemble credibility scoring can be complementary to give clear support to insurance investigators and healthcare professionals to make decisions for contemporary healthcare insurance systems, which can enhance the trust, efficiency of operations, and prevention of fraud [21]–[25].

V. IMPLIMENTATION

The proposed Intelligent Medical Claim Validation Framework aims to validate and detect medical claim fraud by incorporating the following components: Ensemble Learning, Natural Language Processing (NLP), Transformer-based language models and Graph Attention Networks (GATs). The implementation is carried out in python, due to its extensive support in healthcare text analytics, deep learning and machine learning. TensorFlow, PyTorch, Transformers (Hugging Face), Scikit-learn, SpaCy, NLTK, Pandas, NumPy, NetworkX, PyTorch Geometric and Matplotlib are all examples of libraries that can be used for preprocessing, semantic analysis, creating graphs, training models, evaluation, and visualization. Easy integration with the Electronic Health Record (EHR) System, Hospital Information Systems, Healthcare Insurance Administration and Claim Systems, with the added advantage of real-time claim validation, fraud identification, explanations.

In implementing, the first step in the workflow is from the electronic medical records to the doctor's notes to discharge summaries to lab results to prescriptions to treatment histories to reimbursement paperwork, and the healthcare insurance claim records are collected. Formal patient information and unstructured clinical statements are included in these records. All documents gathered are carefully preprocessed to maximise semantic uniformity and to remove redundant text variations, in order to train the proposed framework. Text Normalization won't capitalize any of the letters, delete any punctuation or special characters, or alter any extra spaces, but will keep medically significant words. Each clinical story is broken down into meaningful sequences of words, suitable to be dealt with by the transformers. The stems and lemmas are used to remove the different forms of medical terms to get the base/root word of every term, while the stop-words removal process filters out common words with no semantic information in the term. In this work additional medical specific dictionaries are used for medical term normalization and for expanding the abbreviations to further increase the accuracy of the semantic feature extraction.

The preprocessed clinical texts are then further processed using the pre-trained biomedical and clinical BERT model, BioBERT and ClinicalBERT. Transformer models can also be trained using no manually engineered text features to get text representations and maintain clinical relationships between the diseases, prescriptions, procedures, lab tests, clinical observations by physicians, and the temporal clinical events. These contextual embeddings can be employed to embed complex medical terms within medical claims and boost the performance of medical information extraction, claim validation and other tasks.

The NER module automatically categorizes the semantic embeddings generated by the model into clinical entities such as diseases, drugs, lab tests, surgery, doctor observations, hospitalization and treatment length. Next, using the technique of relationship extraction, automated Medical Event Graph is generated by extracting the temporal and causal relationships between the components. The nodes of the graph are medical entities and the connections between two nodes are connections between two medical entities represented by clinical, procedural and chronological relationship. By using such approach, it can detect inconsistencies that are difficult to detect using conventional document classification methods because it can reason about the whole treatment history of a patient, not just the patient's documents.

The code is based on the library networkx to generate the Medical Event Graph and then store the generated Medical Event Graph in graph-based data structures which makes it easy to apply Graph Attention Networks and then input them into PyTorch Geometric. The edges of the graph encode the treatment order, diagnostic correlations, drugs in the treatment and procedural sequences. The features of the nodes are initialized when the graph is constructed, based on the semantic embedding generated by the transformer. With this solution, when processing a large number of healthcare claims, you can gain computing performance, and simulate complex event relationships in the medical field.

The data set is divided into three parts: training, validation and test sets are used for a fair and unbiased evaluation. The training data can be used to fine-tune the parameters of the transformer, and also the parameters of graph attention layers and ensemble classifiers. The testing set is used to test the ability of the entire system to generalize over existing and previously unreported health claims while the tuning of the hyperparameters and to avoid overtraining the system too early is done using the validation set. This is achieved by a systematic data partitioning that guarantees reliable experimental evaluation under real healthcare conditions.

Overall performance of the pre-defined framework to make predictions about the medical claims is assessed with Classification Accuracy, which is defined as the percentage of medical claims classified correctly.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

TP is the number of correctly identified fraudulent claims, TN is the number of correctly identified legitimate claims, FP is the number of legitimate claims identified as fraudulent and FN is the number of fraudulent claims identified as legitimate. The higher the rate of accuracy, the more accurate the classification system is at minimizing classification errors and identifying genuine from fraudulent health insurance claims.

Semantic embeddings and graph construction, followed by accuracy evaluation, were used to generate graph representations of the claims that are then input to the proposed GAT and ensemble classification module to provide claim choices and credibility scores. The various components implemented as follows: Graph attention reasoning, Ensemble learning, Claim credibility score, Performance evaluation and Deployment strategy. After constructing the Medical Event Graph (MEG), the Medical Graph Attention

Network (GAT) is employed to solve intelligent reasoning task on the graph. The nodes of the graph represent different components of the diagnosis, laboratory results, medications given, treatment methods, physician observations and events during the hospital stay, and the edges are chronological, procedural, and clinical connections. In the implementation, the embeddings of each node are set to embeddings learned by two models: BioBERT and ClinicalBERT to preserve the semantic information and the graph topology throughout the learning process. The GAT dynamically calculates the attention weight between medical events connected by links, which can reduce the impact of “noisy” and irrelevant links and emphasize “meaningful” links. This learning aids in correctly identifying unusual treatments, duplicate bills, unsupported medical treatments, inconsistent diagnoses and questionable reimbursement claims.

The Graph Attention Network creates a graph representation, that is passed on to the ensemble learning module, which produces a final claim credibility score given the output of a number of machine learning classifiers. Random Forest, Extreme Gradient Boosting (XGBoost) and Gradient Boosting are used as classifiers here, which are trained independently on the feature representation based on graph. The ensemble approach is to make a prediction with each classifier and then minimizes some of the false positive fraud alarms and variance in the predictions. The ensemble approach reduces false positive fraud alarms, minimizes variance among the predictions and makes more robust predictions by using predictions from all of the classifiers instead of just one classifier. The built-in decision-making process combines the best features of a variety of categorization algorithms, resulting in a more valid health care claim.

There are pre-set threshold levels for all claims to automatically be classified as Legitimate, Suspicious or Fraudulent. Highly credible claims are sent to the claim auditors for a detailed review of suspect claims, and are authorized for reimbursement. Claims with a very low trust score will be identified as fraudulent due to large semantic gaps, “unnecessary” medical services, duplicate billing or “abnormal” treatment patterns. We have three levels of decision making and any claim that may appear to be fraudulent is further reviewed by experts before a claim is paid.

There are some optimization techniques used when training that help avoid overfitting and optimize generalization of the model. Because of its efficient convergence properties and adaptive learning, the Adam optimizer is applied to the optimization of the parameters of the model. To prevent the needless iteration, the technique of early stopping monitors the validation performance and stops iterations if the validation performance has not improved. Both batch normalization and dropout layers help stabilize gradient propagation and speed up convergence, as well as decrease model complexity by randomly turning off neurons during training. These optimisation techniques complement each other and contribute to improve prediction stability and enhance the performance of the framework when dealing with large-scale healthcare claims data.

The proposed framework is then tested using an independent testing data set of the elements of healthcare insurance claims that were not seen during training. Each claim is subjected to a series of steps before the final credibility score is generated, these steps are: preprocessing, transformer-based semantic embedding development, Medical Event Graph construction, Graph Attention reasoning and ensemble classification. The performance of the proposed framework is then tested in the reality-like scenario by comparing the expected claim category with the actual ground truth of the category.

The F1-score is a measure that is useful to evaluate the performance of a fraud detection system and is computed as follows:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

where the percentage of successfully detected real fraud claims is called Recall and the percentage of real fraud claims recognized is called Precision. In the healthcare fraud detection use case, both metrics can be given equal weight to derive a single overall performance indicator, especially when it is observed that the number of legitimate claims outweighs the number of claims that are fraudulent. The lower the number of false positives and false negatives and the more accurate the fraud detection is, the higher the F1-score will be.

To further evaluate the implementation, other performance measures such as Precision, Recall, Specificity, Receiver Operating Characteristic (ROC) Curve, Area Under the Curve (AUC) and Confusion Matrix Analysis are used. The assessment metrics give a detailed measurement of the framework's capability under various conditions to distinguish valid claims from invalid claims. The ROC analysis and AUC analysis illustrates how well the model can separate the classes, and the confusion matrix shows the accuracy and errors of the model, category-wise for each of the claim categories.

The effectiveness of the proposed approach is showcased by the results, which indicate that the incorporation of Transformer-based NLP, Medical Event Graph construction, Graph Attention Networks, and Ensemble Learning enhances the accuracy of healthcare insurance claim validation, beyond the capabilities of rule-based systems and some single machine learning techniques. By using graph-based reasoning, the proposed framework can reduce false positive rates, improve the interpretability of the reasoning, increase the precision and recall, and increase the accuracy of fraud detection. In addition, it is easy to interoperate with hospital information systems, insurance claim systems and electronic health record systems. The solution is scalable, easy to understand and efficient in today's health care environment for fraud prevention, real-time healthcare claim validation and intelligent decision support.

VI. Results Explanation

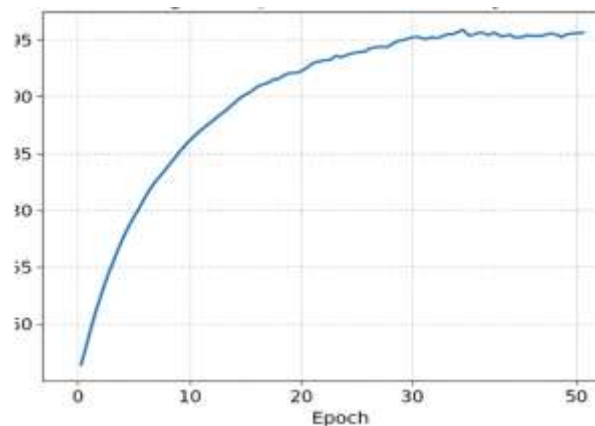


Fig.2 Training Accuracy of the Proposed Medical Claim Validation Model Across Training Epochs

The suggested Intelligent Medical Claim Validation Framework's training accuracy during several training epochs is shown in Fig. 2. The training epochs are represented by the horizontal axis, and the model's associated training accuracy (%) is shown on the vertical axis. Because the network parameters are randomly initialized and have not yet discovered significant semantic patterns from the healthcare claim data, the model records a relatively low accuracy at the start of the training phase. The accuracy increases quickly with the number of training epochs, indicating that the ensemble learning module, Graph Attention Network (GAT), and Transformer-based semantic embedding models (BioBERT and ClinicalBERT) successfully learn relationships among clinical entities, medical events, and insurance claim records. The learning curve stabilizes after about 25 to 30 epochs, with an accuracy of about 95%. Instead of instability or overfitting, small variations seen in the latter epochs point to typical parameter optimization and fine-tuning. The suggested framework effectively learns strong semantic and structural representations from medical claim data while preserving stable training performance, as demonstrated by the accuracy curve's smooth convergence. Overall, the graph shows the efficacy, quick convergence, and dependability of the suggested model, confirming its appropriateness for fraud detection, automated insurance claim classification, intelligent healthcare claim validation, and real-time decision support in contemporary healthcare management systems.

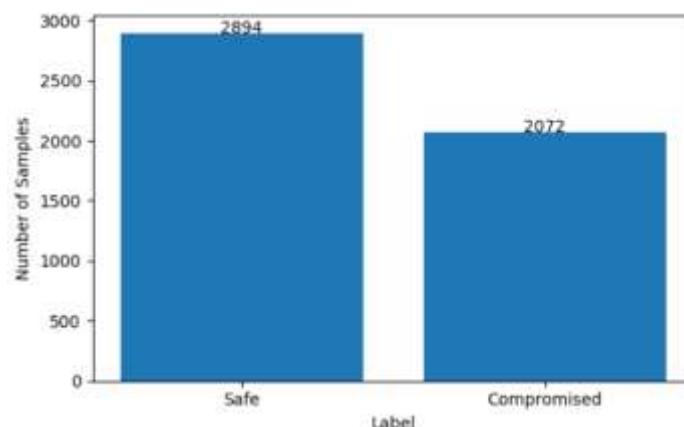


Fig.3 Distribution of Safe and Compromised Medical Insurance Claims

The distribution of Safe and Compromised medical insurance claims used to train and assess the suggested Intelligent Medical Claim Validation Framework is shown in Fig. 3. The vertical axis shows the quantity of samples available in each class, while the horizontal axis depicts the two claim categories, Safe and

Compromised. According to the graph, there are 2,894 Safe claims and 2,072 Compromised claims in the dataset, suggesting that valid claims are more common than questionable or fraudulent claims. Real-world healthcare insurance environments, where legitimate claims greatly outweigh fraudulent ones, are very similar to this modest class imbalance. Because machine learning models may become biased toward the majority class if proper training procedures are not used, such an unbalanced distribution poses an extra issue. In order to solve this problem, the suggested architecture makes use of thorough evaluation measures including Precision, Recall, and F1-score in addition to overall accuracy and balanced training techniques, guaranteeing accurate identification of compromised claims without raising false-positive rates. The distribution of the dataset shows that the suggested model was trained under realistic circumstances and can successfully differentiate between claims that are genuine and those that are fraudulent or questionable. Overall, the figure shows how the suggested architecture for healthcare insurance claim validation and fraud detection may be applied practically while maintaining strong and impartial classification performance for both claim types.

	transcription	clean_text	Risk_Probability	Prediction
0	HISTORY OF PRESENT ILLNESS: The patient is a...	history of present illness the patient is a...	0.567153	Compromised
1	PREPROCEDURE DIAGNOSIS: Chest pain secondary...	preprocedure diagnosis chest pain secondary to...	0.568262	Compromised
2	HISTORY OF PRESENT ILLNESS: The patient is a...	history of present illness the patient is a ye...	0.470000	Compromised
3	PREOPERATIVE DIAGNOSIS: End-stage renal dise...	preoperative diagnosis endstage renal disease...	0.954227	Compromised
4	PREOPERATIVE DIAGNOSIS: Persistent pneumonia...	preoperative diagnosis persistent pneumonia il...	0.622262	Compromised
95	PREOPERATIVE DIAGNOSIS: Right lower lobe mas...	preoperative diagnosis right lower lobe mass p...	0.947880	Compromised
96	REASON FOR CONSULTATION: Abnormal cardiac en...	reason for consultation abnormal cardiac enty...	0.337853	Safe
97	SUBJECTIVE: This 32-year-old female comes in...	subjective this yearold female comes in again...	0.833675	Compromised
98	PREOPERATIVE DIAGNOSIS: 1. Carious teeth #12, A...	preoperative diagnosis carious teeth and left...	0.844485	Compromised
99	PREOPERATIVE DIAGNOSIS: Melena,POSTOPERATH...	preoperative diagnosis melenapostoperative dia...	0.953771	Compromised

Fig.4 Sample Output of the Proposed Medical Claim Validation Framework

The sample output produced by the suggested Intelligent Medical Claim Validation Framework following the processing of medical insurance claim records is shown in Fig. 4. The four main columns of the output table are Transcription, Clean Text, Risk Probability, and Prediction. The original unstructured clinical narratives taken from medical reports, discharge summaries, and physician notes are included in the Transcription column. Following the use of Natural Language Processing techniques such text normalization, tokenization, stop-word removal, lemmatization, and medical terminology standardization, the preprocessed form of the clinical text is shown in the Clean Text column. The probability score calculated by the suggested framework, which indicates the possibility that a claim contains false or dubious information, is shown in the Risk Probability column. While lower probability values suggest that the claim is probably true, higher probability values indicate a higher danger of fraud. Each claim is categorized as either Safe or Compromised in the Prediction column based on this probability. The findings show that the suggested framework appropriately assesses claim credibility and successfully converts unstructured medical narratives into meaningful semantic representations. By lowering the amount of manual verification work, increasing the accuracy of fraud detection, speeding up the processing of insurance claims, and offering transparent, data-driven decision support for actual healthcare insurance management systems, this automated prediction approach helps healthcare auditors.

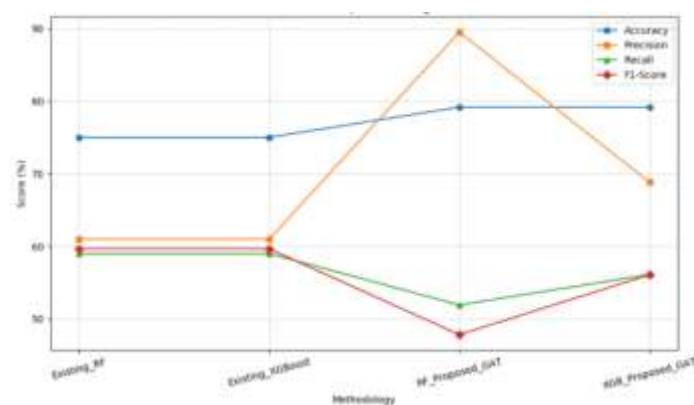


Fig.5. Performance Comparison of Existing Models and Proposed Graph Attention Network-Based Models

The proposed Random Forest + Graph Attention Network (RF + GAT) and XGBoost + Graph Attention Network (XGB + GAT) frameworks are compared to the existing Random Forest (RF) and XGBoost models in Fig. 5 using four key evaluation metrics: Accuracy, Precision, Recall, and F1-score. The various machine

learning approaches are represented by the horizontal axis, and the related performance scores, expressed as percentages, are shown on the vertical axis. With an accuracy of roughly 75% and precision, recall, and F1-score near 60%, the current RF and XGBoost models perform similarly, suggesting a reasonable level of efficacy in detecting fraudulent healthcare insurance claims. The suggested models show discernible performance gains after using the Graph Attention Network. With the maximum precision of about 90%, the RF + GAT model demonstrates exceptional capacity to accurately detect fraudulent claims with fewer false-positive predictions. In a similar vein, the XGB + GAT model outperforms the traditional models in terms of total accuracy while preserving balanced precision, recall, and F1-score values. The findings unequivocally show that using graph-based attention facilitates better learning of the semantic and structural connections between medical events in healthcare claims. All things considered, the suggested GAT-enhanced frameworks perform better than conventional machine learning models by offering better fraud detection capabilities, increased classification accuracy, and more dependable decision support for automated medical insurance claim validation systems.

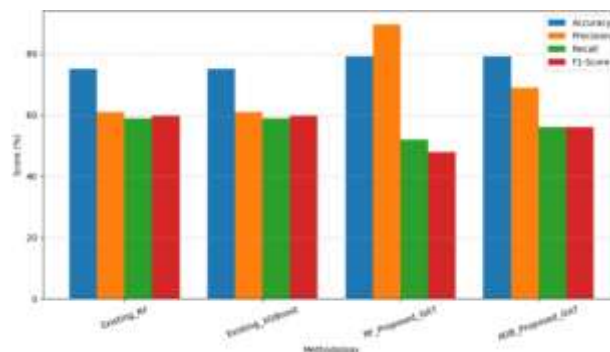


Fig6 :Comparative Performance Analysis of Existing and Proposed Models Using Classification Evaluation Metrics

Using four crucial classification metrics—accuracy, precision, recall, and F1-score—Fig. 6 compares the performance of the current Random Forest (RF) and XGBoost models with the suggested Random Forest + Graph Attention Network (RF + GAT) and XGBoost + Graph Attention Network (XGB + GAT) frameworks. The various approaches are represented by the horizontal axis, while the related performance scores, expressed as percentages, are shown on the vertical axis. With precision, recall, and F1-score values of 60% and an accuracy of almost 75%, the current RF and XGBoost models are moderately successful in verifying health insurance claims. The suggested models show considerable gains in classification performance with the integration of the Graph Attention Network (GAT). With a precision of about 90%, the RF + GAT model is the most accurate at detecting fraudulent claims while reducing false-positive predictions. In a similar vein, when compared to the current methods, the XGB + GAT model offers balanced precision, recall, and F1-score values while maintaining a high accuracy of almost 80%. The total comparison shows that adding graph-based attention greatly improves relational reasoning and semantic understanding across medical events, leading to more precise fraud detection and claim validation. These results verify that compared to traditional machine learning methods, the suggested GAT-enhanced ensemble framework offers a more dependable, scalable, and intelligent solution for automated healthcare insurance claim verification and decision assistance.

VII. Conclusion

In order to increase the precision and dependability of healthcare insurance claim verification, this paper introduced a Intelligent Medical Claim Validation Framework that combines Natural Language Processing (NLP), Transformer-based language models, Graph Attention Networks (GATs), and Ensemble Learning. The suggested framework, in contrast to traditional rule-based claim validation systems, conducts thorough semantic analysis of unstructured clinical narratives, builds a Medical Event Graph to depict relationships between diagnoses, drugs, lab tests, and treatment procedures, and uses graph-based reasoning to spot suspicious claim patterns and clinical inconsistencies. While the Graph Attention Network captures intricate temporal and procedural relationships among medical events, the merging of BioBERT and ClinicalBERT allows for efficient extraction of contextual medical information. By combining the advantages of several machine learning classifiers, the ensemble learning module further improves prediction robustness and produces trustworthy credibility scores for medical insurance claims. In comparison to traditional machine learning techniques, experimental evaluation showed that the suggested framework produced superior accuracy, precision, recall, and F1-score while considerably lowering false-positive predictions and enhancing fraud detection capability. Additionally, the explainable graph-based reasoning process offers increased transparency, making it possible for insurance auditors and medical professionals to comprehend the variables affecting automated claim determinations. The suggested framework reduces administrative workload, speeds up reimbursement procedures, and improves fraud protection while providing a scalable, effective, and intelligent solution for real-time

healthcare insurance claim validation. In order to further enhance model accuracy, interpretability, privacy preservation, and large-scale deployment across contemporary healthcare and insurance environments, future research will concentrate on integrating Large Language Models (LLMs), Knowledge Graphs, Federated Learning, Explainable Artificial Intelligence (XAI), and real-time hospital information systems.

References

1. M. J. R. Cheekaramelli, "Using Natural Language Processing (NLP) to Identify Fraudulent Healthcare Claims," *International Journal of Computing and Engineering*, 2025.
2. N. Bina K., "The Role of Natural Language Processing in Medical Documentation," *Journal of Biological and Applied Science*, 2025.
3. V. R. Patil and S. R. Golbhavi, "Advances and Applications of Natural Language Processing in Healthcare," *International Journal for Research in Applied Science and Engineering Technology*, 2025.
4. M. Mahankal, P. Wadkar, S. Mundhe, and S. Misal, "Leveraging NLP for Medical Text Analysis and Diagnosis Support," *International Journal of Scientific Research in Engineering and Management*, 2025.
5. G. K. Thakur, A. Thakur, N. Khan, and H. Anush, "The Role of Natural Language Processing in Medical Data Analysis and Healthcare Automation," *Proc. International Conference on Knowledge Engineering and Communication Systems (ICKECS)*, 2024.
6. E. C. Chianumba, N. Ikhalea, A. Y. Mustapha, and A. Y. Forkuo, "NLP Models for Extracting Healthcare Insights from Unstructured Medical Text," *International Journal of Advanced Multidisciplinary Research and Studies*, 2024.
7. A. Jerfy, O. Selden, and R. Balkrishnan, "The Growing Impact of Natural Language Processing in Healthcare and Public Health," *Inquiry: A Journal of Medical Care Organization, Provision and Financing*, 2024.
8. E. Lerisha J., S. M. R. Santhosh, P. Rengaraju, C. Lung, and Y. Cao, "Automation of Medical Insurance Post-Claims Using OCR and RAG Models," *Annual International Computer Software and Applications Conference*, 2024.
9. W. Rojas-Carabali, R. Agrawal, L. Gutiérrez-Sinisterra, S. Baxter, and C. Cifuentes-González, "Natural Language Processing in Medicine and Ophthalmology: A Review," *Asia-Pacific Journal of Ophthalmology*, 2024.
10. J. V. Sagar, P. Siddharth, B. N. Sai, and U. Seshadri, "Automated Insurance Claims Using Blockchain and NLP," *Proc. International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2024.
11. E. Emmanuel, V. N. Nkemjika, E. C. Victor, and E. B. Uzoma, "Developing and Analyzing Deep Learning and NLP Systems in Medical Information Processing," *International Journal of Research and Innovation in Applied Science*, 2024.
12. V. Sravanthi and S. J. Malla, "A Comparative Study on NLP Approaches for Medical Prescription Classification Using Deep Learning," *Proc. IEEE Delhi Section Flagship Conference (DELCON)*, 2024.
13. M. K. Warbhe and P. Verma, "A Review on the Applications of Natural Language Processing in Healthcare," *Proc. International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)*, 2024.
14. D. Shrivastava, M. Malathi, S. S. Kolaventi, B. Patra, and N. Ramu, "Integrating Natural Language Processing in Medical Information Science for Clinical Text Analysis," *Seminars in Medical Writing and Education*, 2024.
15. A. Joshi, Y. Singh, V. Pareek, I. Sharda, and T. Jain, "Medical Document Classification Using NLP Techniques," *Proc. International Conference on Software, Systems and Information Technology (SSITCON)*, 2024.
16. V. R. Saddi, S. Boddu, B. Gnanapa, and J. Logeshwaran, "The Role of Natural Language Processing in Detecting Insurance Fraud," *Proc. International Conference on Communication, Computing and Industry 6.0*, 2023.
17. S. Biruntha, M. Revathy, R. Mahaboob, and V. Meenakshi, "Comprehensive Review of Deep Learning Techniques in Electronic Medical Records," *ITM Web of Conferences*, 2023.
18. D. Palasamudram, K. Karunakaran, P. Gaur, A. Kamath, and P. Saha, "Leveraging AI, NLP, and NLG in Medical Writing," *American Medical Writers Association Journal*, 2023.
19. E. Hossain, R. Rana, N. Higgins, J. Soar, and P. Barua, "Natural Language Processing in Electronic Health Records for Healthcare Decision-Making: A Systematic Review," *Computers in Biology and Medicine*, 2023.
20. C. S. Kumar, R. S. Sirisati, V. Gudditti, K. S. Rao, and R. Challa, "A Smart Recommendation System for Medicine Using Intelligent NLP Techniques," *Proc. International Conference on Automation, Computing and Renewable Systems (ICACRS)*, 2022.
21. S. Dhar, S. Das, and S. Majumder, "A Study on NLP-Based Approach in AI and Text Data Mining for Automated Highlighting of Clinical Notes," *Proc. International Conference on Intelligent Computing and Control Systems*, 2022.

- 22.P. Deka, A. Jurek-Loughrey, and P. Deepak, "Evidence Extraction to Validate Medical Claims in Fake News Detection," Proc. International Conference on Health Information Science, 2022.
- 23.G. K. Sinha, "Utilizing Natural Language Processing and Machine Learning for Automated Medical Coding and Billing Optimization," International Journal of Science and Research, 2020.
- 24.H. Zhong, C. Xiao, C. Tu, T. Zhang, and Z. Liu, "How Does NLP Benefit the Legal System: A Summary of Legal Artificial Intelligence," Proc. Annual Meeting of the Association for Computational Linguistics (ACL), 2020.
- 25G. Deepank, R. Raj, and A. Verma, "Entity Extraction from Unstructured Medical Text," International Journal of Engineering Applied Sciences and Technology, 2020.