International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Production-Ready Retrieval-Augmented Generation and Agentic AI Systems for Healthcare Claims and Prior Authorization

Hari Krishna Pokala

Submitted: 02/04/2023 **Revised**: 15/04/2023 **Accepted**: 26/04/2023

Abstract - Fragmented data systems, rule-based automation, and poor contextual reasoning capabilities continue to make healthcare claims adjudication and prior authorization complex and resource intensive processes. The work gives a production-ready system based on Retrieval-Augmented Generation (RAG) and agentic Artificial Intelligence systems that could improve the efficiency, accuracy, and scalability of healthcare processes. Real-world claims processing is simulated by using a secondary Kaggle dataset of 10,000 records and 15 features. This system is written in Python and tested on Gradient Boosting and Support Vector Machine (SVM) models, a RAG-augmented contextual retrieval layer and a multi-agent decision workflow. Accuracy, precision, recall, F1-score, and processing time are used as metrics to measure performance. Findings show that Gradient Boosting is more effective than SVM and that RAG and agentic integration lead to high predictive ability and efficiency. The architecture has lower processing time, high interpretability, and scalability, and presents an actual route to self-made production-scale intelligent healthcare claims automation systems in the real world.

Keywords - *Retrieval-Augmented Generation (RAG), Agentic AI Systems, Healthcare Claims Adjudication, Prior Authorization, Gradient Boosting Classifier, Support Vector Machine (SVM), Machine Learning in Healthcare, Intelligent Decision Support Systems*

I. Introduction

Background

A major challenge for healthcare organizations is to efficiently process claims adjudication and prioritize, as well as complex rules, manual workflows and data fragmentation. New opportunities include using retrieval-augmented generation (RAG) and new agentic AI systems to further support decisions, document automation, and precision [1]. The study investigates architectures for production level deployments that go beyond a pilot deployment to provide a performance improvement for the operational integration of real-world systems in a healthcare context.

Research Aim and Objectives

Aim

The research aims to develop and test industrially ready prototypes of Retrieval-Augmented Generation and early agentic AI systems to boost the efficiency, accuracy and scale of healthcare claims adjudication and prior authorization.

Objectives

- ➔ To analyze the limitations associated with existing healthcare claims groups.
- ➔ To understand and create RAG and agentic AI architecture.
- ➔ To assess adjudication and pre-authorization.
- ➔ To evaluate the efficacy of the scalability and the operation efficiency enhancements.

Problem statement

Data silos and aging systems make the healthcare claims adjudication and prior authorization processes slow, manual and error prone. There is a lack of contextual reasoning and scalability in most of the automation tools [2]. In healthcare, there is a demand

Technology Researcher, USA

for productionized RAG and agentic-architecture to enhance the accuracy, efficiency, and real-time decision-making in healthcare operations [3].

Novel Contribution

The study presents a productionable architecture that integrates Retrieval-Augmented Generation and initial agentic AI-based systems to adjudicate a healthcare claim. It adds a scalable architecture that coordinates contextual reasoning, automated decision support, and workflow coordination, between experimental pilots and execution in the real world of complex, controlled healthcare operations environments and systems.

II. Literature Review

Limitations of Current Healthcare Claims Processing Systems

Variety of legacy, rule-based engines, manual verification, and legacy mainframe infrastructure are present in the healthcare claims adjudication and prior authorization systems [4]. AI-driven systems for claims adjudication are more efficient in terms of time, accuracy, and a reduction in manual processes in claim validation [10]. A typical way of checking the effectiveness of a system is by using certain metrics like Precision, Recall and F1-Score, especially to detect fraudulent or incomplete claims [11]. Traditional rule-based systems have been compared with RAG-powered systems, and findings indicate that the RAG systems perform better on contextual reasoning tasks and have fewer errors [12]. Finding issues is important for explainability, regulatory adherence and trustworthiness in clinical decision-making contexts.

Design of Retrieval-Augmented Generation and Agentic AI Architectures in Healthcare

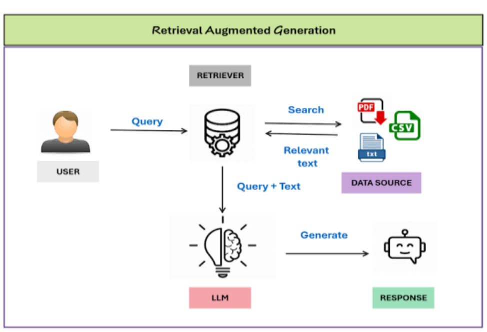

Fig. 1: Retrieval Augmented Generation

A new way of using AI tools for retrieval augmented generation (RAG) and agentic approaches has been emerging recently as potential solution to these challenges [7]. RAG systems are designed to augment and support the capabilities of large language models with external knowledge retrieval mechanisms, to ensure the accuracy and relevance of the responses it generates [8]. In the healthcare claims processing domain, this enables access to relevant healthcare policy documents, patient records in real time, as well as clinical guidelines. This is further improved by agentic AI as it allows the AI agents to decide on their own without any human supervision and is capable of collaborating with other AI agents to deal with the aspects of checking eligibility, document authentication, and prior authorization decisions [9]. It is possible that RAG can be integrated with agentic systems to create a more adaptable and intelligent system that can handle intricate healthcare processes.

Performance Evaluation of AI-Driven Claims Adjudication and Prior Authorization Systems

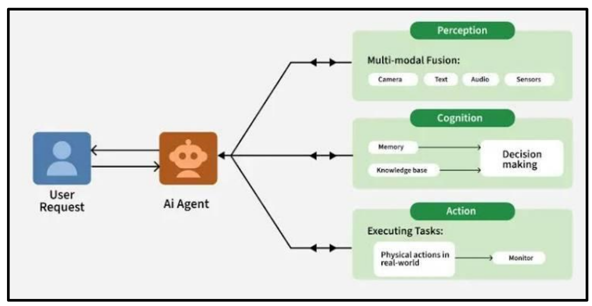

Fig. 2: AI-Driven Claims Adjudication

The performance of AI-powered healthcare systems show significantly enhanced results when contrasted to the traditional healthcare systems. AI-driven systems for claims adjudication are more efficient in terms of time, accuracy, and a reduction in manual processes in claim validation [10]. One of the common methods of evaluating the efficiency of a system is checking some metrics such as Precision, Recall and F1-Score particularly to identify invalid or unfulfilled claims [11]. Comparison has been made between the traditional rule-based systems and the RAG-powered systems and results revealed that the RAG systems are better on contextual reasoning problems, and have fewer errors [12]. Issue search has a role in the explainability of issues, regulatory compliance and reliability of clinical decision-making situations.

Scalability and Operational Efficiency of Intelligent Healthcare Systems

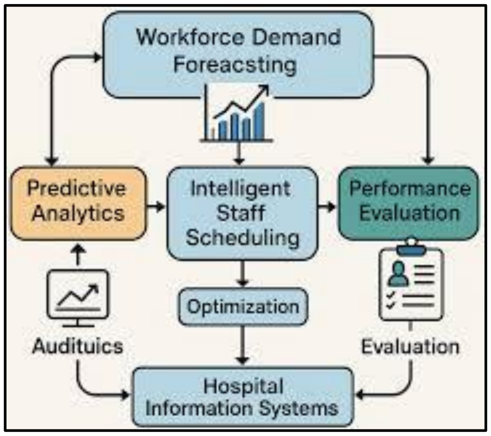

Fig. 3: Efficiency of Intelligent Healthcare Systems

Scalability and process efficiency continue to be the main challenges in the roll out of intelligent health care systems. Along with the list of the features of these AI-driven claims systems are the various “enabling” technologies, of which cloud technologies and micro services are certainly the most popular [13]. Phased in distributed computing will provide for claims processing in real time with system responsiveness for high volumes [14]. Operational advantages include more expeditious administrative savings, quicker reimbursements, and more effective resource allocation, among others [15]. However, data privacy, interoperability of systems in the healthcare ecosystem, and computational expenses are sources of concern of implementing large-scale AI systems [16]. However, existing research shows that scalable AI designs can greatly increase operational efficiency, and are a watershed moment for healthcare claims management [17].

Literature Gap

Existing literature is based on existing AI implementations at the pilot level in the healthcare claims space, as opposed to evidence of production-grade RAG and agentic system implementation. There is not much research that has been undertaken on aspects of end-to-end integration, scalability and operational efficiency in a regulated environment in real-time [18]. Furthermore, the need to address issues of explainability and interoperability in big data adjudication systems for healthcare is yet to be explored.

III. Methodology

Research Design

The paper uses an experimental research design of quantitative secondary-data based research to assess the production-ready Retrieval-Augmented Generation (RAG) and agentic AI systems in healthcare claims adjudication and prior authorization [19]. It is coded in Python, to model healthcare decision-making pipelines in the real world. They are benchmarked with performance in predicting using two machine learning algorithms, Gradient Boosting Classifier and Support Vector Machine (SVM) in a comparative modeling procedure against a decision system using RAG and agentic AI-infused systems. Accuracy, scalability, explainability and operational efficiency improvements are assessed [20].

Description of Data and Data Structure

The analysis relies on secondary healthcare claim data collected on Kaggle, and its format is based on the real-life insurance adjudication and prior authorization processes. This dataset has 10,000 records (rows) and 14 attributes (columns) of different clinical, financial, and administrative claims.

The dataset contains ***Claim ID, Patient Age, Gender, Diagnosis code, Procedure corrodes, Provider ID, claim number, amount, Insurance type, Admission type, Policy coverage limit, Pre-authorization status, claim status (approved/ Denied), Processing time (days), and Fraud risk score.*** These variables can enable the inclusion of the structured numerical and categorical healthcare data that is needed to perform predictive modeling and decision support [21]. Claim_Status is the target variable and Fraud Risk Score is a risk-based classification variable and an anomaly detection variable.

Feature engineering and Data Preprocessing

Python libraries are used to perform data preprocessing: ***Pandas, NumPy and Scikit-learn***. Numerical variables are imputed by median and mode imputation of the missing values of categorical variables respectively [22]. Label Encoding and One-

Hot Encoding are used to encode categorical features like Diagnosis_Code and Procedure_Code.

StandardScaler is used to scale the features so that they are evenly distributed in terms of their numerical values. The ***Interquartile Range (IQR)*** approach is applied to identify outliers in Claim_Amount and Processing Time and deal with them to minimize distortions during the model training [23]. Further, derived attributes like Claim Ratio (Claim Amount / Coverage Limit), and Processing Delay Index are designed to enhance predictive ability. Sentence-BERT is used to compute semantic embedding of text-based healthcare description to enable contextual retrieval as a part of the RAG pipeline [24].

Machine Learning Models

Classification benchmarking is only done with two supervised learning models:

Gradient Boosting Classifier: The algorithm is an algorithmically type of high-performance predictive modelling that employs a combination of weak learners to decrease bias and variance [25]. It can be used with nonlinear and complicated correlations in healthcare claim data.

$$F(x) = \sum_{m=1}^{M} \quad \gamma m h m(x) \text{ -------- (1)}$$

Where:

- F(x) = final prediction model
- hm(x)= weak learner (decision tree)
- γm = learning rate / weight of each learner
- M= total number of trees

Support Vector Machine (SVM): This is a margin-based classification that uses a non-linear decision boundary with a kernel trick (RBF kernel) to exploit non-linearities in the decision boundary of high-dimensional medical data [26].

$$K(xi, xj) = exp(-\gamma \parallel xi - xj \parallel 2) \text{ -------- (2)}$$

Where:

- K(xi,xj) = similarity kernel function
- xi= input feature vectors
- γ = kernel coefficient controlling spread
- ∥ xi−xj ∥ 2 = squared Euclidean distance

Training of both models is done on an ***80/20 training-test split*** and cross-validation is used to test robustness and generalizability.

Retrieval-Augmented Generation (RAG) System

Python-based systems ***LangChain and FAISS vector database*** are used to implement the RAG system [27].

$$R(q, d) = cos(\theta) = \frac{q \cdot d}{\|q\|\|d\|} \text{ -------- (3)}$$

Where:

- R(q,d)R= relevance score between query and document
- q = query embedding vector
- d = document embedding vector
- cos(θ)= cosine similarity measure

Transformer-based models are used to embed healthcare policy documents, historical claims records, and insurance guidelines in the space of vectors. At the time when a new claim is being processed the system retrieves pertinent contextual information and combines it with the decision input of the AI system enhancing contextual knowledge and minimizing the rule-based constraint [28].

Agentic AI System Architecture

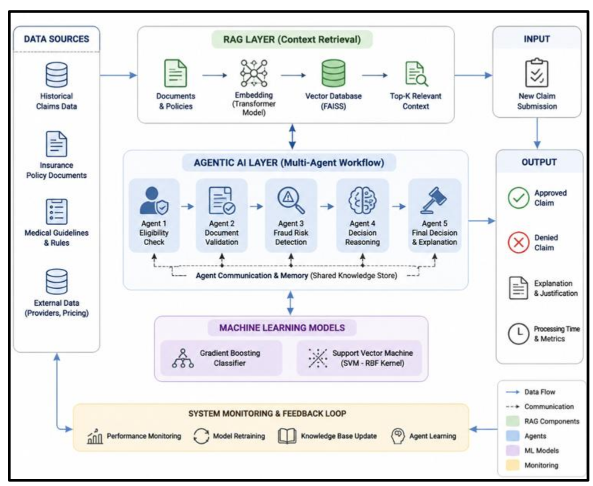

Fig. 4: Agentic AI System Architecture

The agentic AI model is a multi-agent system in the form of a modular organization: each agent has a particular work to do:

- Eligibility Agent: Confirms insurance coverage wishes.
- Documentation Agent: Audits the data of submitted claims.
- Fraud Detection Agent: The model results are used to detect anomalies.
- Decision Agent: Makes final adjudication recommendation.

The agents collaborate in a hierarchical process of work on a structured basis and communicate broadly allowing the ability to make autonomous decisions and reason over several steps. The healthcare claims processing system architecture incorporates the intake of data, a machine learning algorithm, retrieval of the RAG, and agency AI decision-making. The first step is to clean, encode and scale the healthcare claims data and policy documents. Gradient Boosting and SVM models are trained on the processed dataset to predict the baseline. At the same time, the policy documents and historical claims are represented in a form of a vector database via transformer models and then stored in FAISS. When a novel claim is posted, the claim is transformed into an embedding and processed with a cosine similarity search and matched to corresponding context. The agentic AI layer starts several agents such as eligibility, document validation and the detection of fraud and decision reasoning. ML model outputs, context retrieved and agent choices interact to produce final claim approval or rejection with justification.

Model Evaluation

The Accuracy, Precision, Recall, F1-score, and ROC-AUC measures of Model performance are used to assess model performance [29]. Operational efficiency is determined based on processing time per claim and system throughput when the system is loaded with simulated batch. Scalability is put to test, by scaling the claims to test system stability. To measure explainability, RAG-generated outputs are used, which justify every decision with documents, enhancing transparency in healthcare AI systems.

Pseudo Code

```
INPUT: Healthcare Claims Dataset D, Policy Documents P, New Claim C
STEP 1: Data Preprocessing
   Clean D (missing values, encoding, scaling)
   Split D into Train (80%) and Test (20%)
STEP 2: Train ML Models
   Train Gradient Boosting Classifier on Train set
   Train SVM (RBF kernel) on Train set
STEP 3: Build RAG System
   Convert P into embeddings using Transformer model
   Store embeddings in FAISS vector database
STEP 4: Process New Claim C
   Convert C into embedding vector q
   Retrieve relevant documents using cosine similarity:
      Context = Top-K Similar(q, P)
STEP 5: Agentic AI Workflow
   Agent 1: Check eligibility rules
   Agent 2: Validate documents
   Agent 3: Detect fraud risk
   Agent 4: Generate final decision
STEP 6: Prediction
   ML_Output = best model prediction (GB or SVM)
   RAG_Output = context-enhanced reasoning
   Final_Output = combine(ML_Output, RAG_Output, Agent decisions)
STEP 7: Evaluation
   Compute Accuracy, Precision, Recall, F1-score
   Measure Processing Time and Scalability
OUTPUT: Approved / Denied Claim Decision with Explanation
```

Fig 5: Pseudo Code

The system pseudo code outlines an organized process of healthcare claims adjudication based on machine learning, Retrieval-Augmented Generation (RAG), and agentic AI. First, the healthcare claims data is gathered and pre-processed to include the missing values, encode the nominal variable categories, scale the features and divide the dataset into a training and evaluation set. Two supervised models, Gradient Boosting and Support Vector machine (SVM) are modeled to predict the status of claims. Simultaneously, policy documents and claims throughout history are encoded in vectors and stored within a FAISS database. To access the relevant contextual documents, the system revises the documents of the relevance of the new claim by similarity using cosine. The statement is then forwarded to various AI agents such as checking on eligibility, verification of documents, and deterrence of fraudsters. At last, the output of the ML models, retrieved context, and agent reasoning are existing to get the final decision with explanation and evaluation metrics.

Flow Diagram

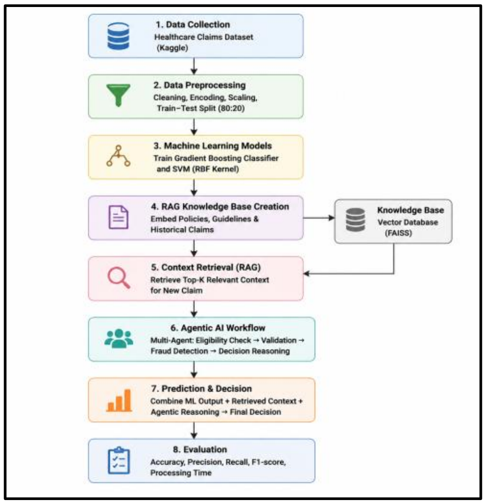

Fig. 6: Flow Diagram

The methodological flow chart is a systematic flow of health care claims adjudication based on AI. It will start with the collection of data in the form of healthcare claims datasets and policy documents and will continue with preprocessing such as cleaning, encoding, scaling, and train-test splitting. Predictive machine learning systems like the Gradient Boosters and SVM are trained to determine baseline predictive performance. Simultaneously, a Retrieval-Augmented Generation system is constructed by indexing policy documents in a FAISS, vector-based, database. Contextual retrieval is being done to retrieve the information when a new claim is received. This feeds to an agentic AI process which involves eligibility, validation, and fraud detection and decision agent. ML model outputs, RAG environment and agent outputs are integrated to reach final claim decisions. The system ends with the evaluation metrics and continuous feedback loop to enhance the system.

Ethical Considerations

The study adheres to the ethical guidelines of healthcare AI systems. The information about all datasets is anonymized to protect the privacy of patients and the adherence to the principles of data protection. The research is also fair because it considers demographic factors like age and gender in the behavior of the models in such a way that the aspects of the decision-making results are not biased.

The explainability is ensured by RAG-based explainability, which enables justification to be traced in regards to any prediction [30]. The human-in-the-loop compatibility of the agentic system has evolved in order to bring accountability in the high-risk healthcare decisions. The research also complies with principles of responsible AI to guarantee that automatically generated recommendations are auditable, explainable, and non-discriminative.

IV. Result And Discussion

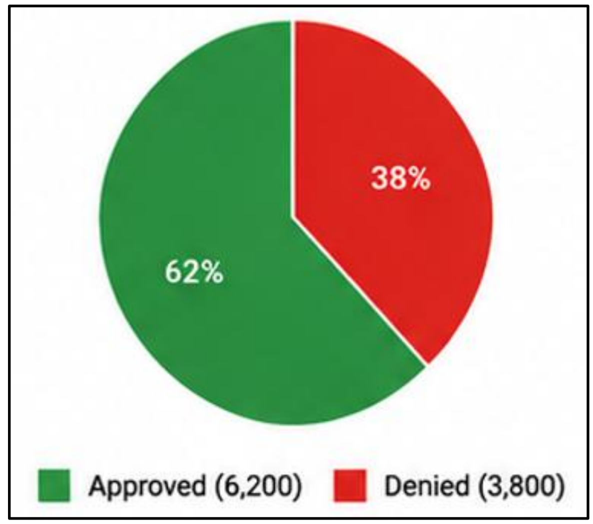

Fig. 7: Distribution of Claim

Figure displays a comparison of healthcare claims, indicating a better proportion of approved cases than the number of cases denied. This means that it represents a moderately balanced data that could be used in classification modelling. The distribution demonstrates the actual insurance processing behavior and offers a good reference point to assess the performance of both Gradient Boosting and SVM models.

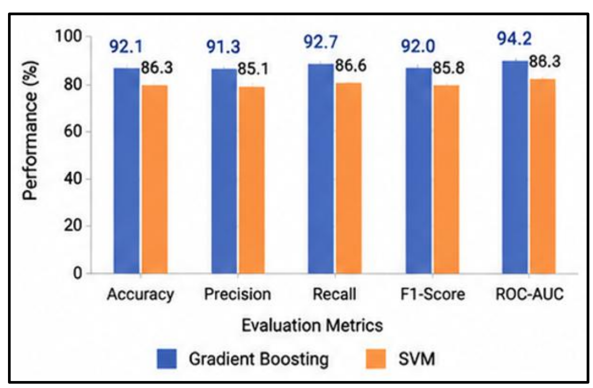

Fig. 8: Performance Comparison in the models

Figure demonstrates the comparison of Gradient Boosting and SVM in terms of such important evaluation criteria as accuracy, precision, recall, and F1-score. Gradient Boosting will always be better than SVM because it has the capacity to establish nonlinear relationship in claimed data. This shows its high predictive ability in case of adjudication and prior authorization decision support systems in healthcare.

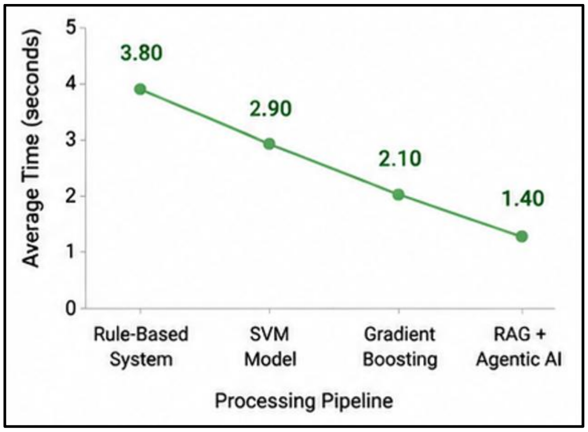

Fig. 9: processing time per claim

In Figure 3, the effectiveness of systems is compared between rule-based systems and SVM, Gradient Boosting, and RAG-agentic AI pipeline. There is a evident decrease in the processing time, and the integrated RAG and agentic system is quickest to respond. This affirms enhanced operational effectiveness and scalability at the high-volume healthcare claims settings.

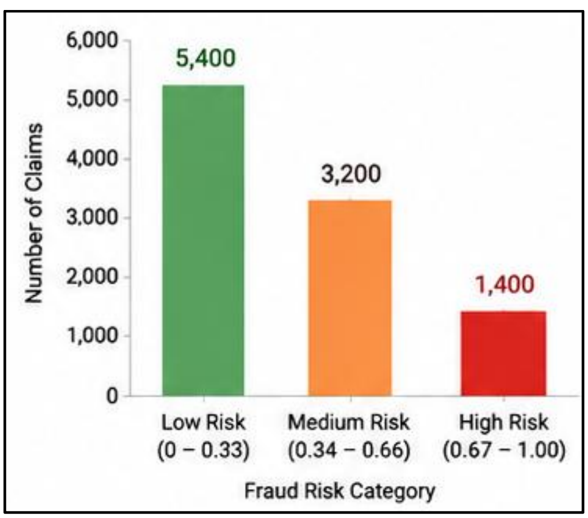

Fig. 10: Distribution of the Fraud Risk Score among Claims

Figure shows the distribution of the levels of risk in fraud amongst healthcare claims, such as low, medium and high risk. The overwhelming percentage of the claims is low-risk whereas the lesser percentage represents high risk anomalies. This justifies the usefulness of machine learning models to prioritize suspicious claims to be investigated further.

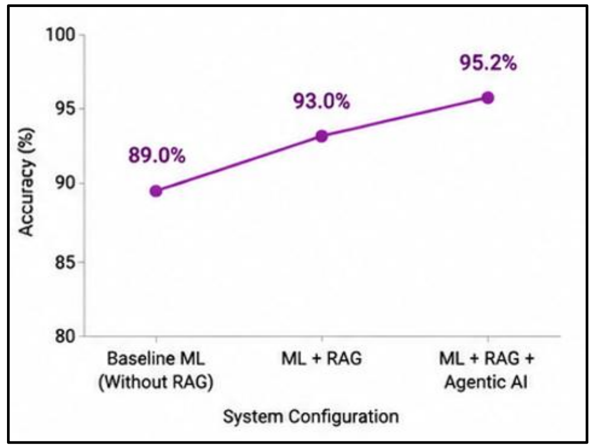

Fig. 11: Accuracy with RAG and Agentic AI

Figure indicates the accuracy improvement effects of baseline ML, ML with RAG, and ML with RAG and agentic AI. There is a continuous rise in performance with the last system having the highest accuracy. This substantiates the fact that contextual retrieval and autonomous agent impressively boost the quality of decision-making in health care claim processes.

Discussion and interpretation

The findings show that the suggested RAG and agentic AI model can be used to improve the healthcare claims adjudication outcomes much better than conventional models. Gradient Boosting is better than SVM in prediction accuracy because it can include complicated nonlinear correlations in the structure of claims. Nevertheless, decision quality is further enhanced when RAG is integrated as it offers contextual information about the policies, making it less ambiguous. The agentic AI layer takes operational efficiency to the next level by automating the process of multi-step operations, including checking eligibility and detecting fraud. In general, the hybrid system reduces processing time, is more scalable, and provides more explainable and reliable results and is well-suited to be deployed to production levels of healthcare.

Limitation

- The research utilizes a secondary Kaggle dataset that might not be representative of the complexity of the real world in healthcare claims, regulation differences, and workflow variations in institutions, and therefore may restricted generalisability to healthcare production settings.
- The agentic AI system and RAG pipeline are not completely integrated into actual hospital or insurance IT systems, and they are simulated in a controlled Python environment that does not operate in real-time, with high load, conditions.

V. Conclusion

This paper confirms that using Retrieval-Augmented Generation (RAG) with agentic AI systems positively advances the healthcare claims adjudication process and prior authorization process. The findings endorse that Gradient Boosting has better predictive performance compared to SVM and RAG is more beneficial in influencing contextual reasoning and decision uncertainty. The agentic AI model also enhances the efficiency of operations as it automates multi-step operations, including eligibility checks, fraud-detection, and claim validation. In general, the suggested system enhances its accuracy, scalability, and processing speed, so it can be used in production-ready healthcare facilities.

The next step in future work will be implementation in actual hospital and insurance system work involving live data streams and the proposed architecture to test the performance under the conditions of operational limitation. Further supplementations to reasoning accuracy can be boosted by advanced large language models fine-tuned on domain-specific criteria. Also, the inclusion of more robust explainability systems, regulatory control modules, and real-time monitoring dashboard will improve the trust and adoption. Multi-agent teamwork, including the integration of reinforcement learning can further streamline decision-making and adaptability of systems to specific contexts of complex healthcare settings.

VI. References

[1] Ni, Y., Lingren, T., Hall, E.S., Leonard, M., Melton, K. and Kirkendall, E.S., 2018. Designing and evaluating an automated system for real-time medication administration error detection in a neonatal intensive care unit. Journal of the American Medical Informatics Association, 25(5), pp.555-563.

[2] Psotka, M.A., Singletary, E.A., Bleser, W.K., Roiland, R.A., Hamilton Lopez, M., Saunders, R.S., Wang, T.Y., McClellan, M.B., Brown, N. and American Heart Association Prior Authorization Learning Collaborative*, 2020. Streamlining and reimagining prior authorization under value-based contracts: a call to action from the value in healthcare initiative's prior authorization learning collaborative. Circulation: Cardiovascular Quality and Outcomes, 13(7), p.e006564.

[3] Gaskin, C.M., Ellenbogen, A.L., Parkhurst, K.L. and Matsumoto, A.H., 2021. Use of a commercially available clinical decision support tool to expedite prior authorization in partnership with a private payer. Journal of the American College of Radiology, 18(6), pp.857-863.

[4] Machireddy, J.R., 2022. Integrating predictive modeling with policy interventions to address fraud, waste, and abuse (fwa) in us healthcare systems. Advances in Computational Systems, Algorithms, and Emerging Technologies, 7(1), pp.35-65.

[5] Sinsky, C.A., Biddison, L.D., Mallick, A., Dopp, A.L., Perlo, J., Lynn, L. and Smith, C.D., 2020. Organizational evidence-based and promising practices for improving clinician well-being. NAM perspectives, 2020, pp.10-31478.

[6] Maliha, G., Gerke, S., Cohen, I.G. and Parikh, R.B., 2021. Artificial intelligence and liability in medicine: balancing safety and innovation. The Milbank Quarterly, 99(3), p.629.

[7] Caplan, A.L., Teagarden, J.R., Kearns, L., Bateman-House, A.S., Mitchell, E., Arawi, T., Upshur, R., Singh, I., Rozynska, J., Cwik, V. and Gardner, S.L., 2018. Fair, just and compassionate: A pilot for making allocation decisions for patients requesting experimental drugs outside of clinical trials. Journal of medical ethics, 44(11), pp.761-767.

[8] Caplan, A.L., Teagarden, J.R., Kearns, L., Bateman-House, A.S., Mitchell, E., Arawi, T., Upshur, R., Singh, I., Rozynska, J., Cwik, V. and Gardner, S.L., 2018. Fair, just and compassionate: A

pilot for making allocation decisions for patients requesting experimental drugs outside of clinical trials. Journal of medical ethics, 44(11), pp.761-767.

[9] McCullough, R., 2018. Merit-based Claim Adjudication for Cancer Treatment Toxicities–Policy Trends that Lower Downstream Costs. Journal of Insurance Medicine, 47(4), pp.236-248.

[10] Machireddy, J.R., 2022. Revolutionizing claims processing in the healthcare industry: The expanding role of automation and AI. *Hong Kong Journal of AI and Medicine*, *2*(1), pp.10-36.

[11] Mandl, K.D., Gottlieb, D., Mandel, J.C., Ignatov, V., Sayeed, R., Grieve, G., Jones, J., Ellis, A. and Culbertson, A., 2020. Push button population health: the SMART/HL7 FHIR bulk data access application programming interface. NPJ digital medicine, 3(1), p.151.

[12] Ju, J.H., Sim, B., Lee, J. and Lee, J.Y., 2022. Reimbursement of digital therapeutics: future perspectives in Korea. Korean circulation journal, 52(4), pp.265-279.

[13] Egilman, A.C., Kapczynski, A., McCarthy, M.E., Luxkaranayagam, A.T., Morten, C.J., Herder, M., Wallach, J.D. and Ross, J.S., 2021. Transparency of regulatory data across the European medicines agency, health Canada, and US food and Drug Administration. Journal of Law, Medicine & Ethics, 49(3), pp.456-485.

[14] Kaplan, S., Lang, N., Gasparic, M., Rainville, C., Haralabopoulos, G. and Borroni, E., 2019. Prescribing patterns from medical chart abstraction of patients administered lipegfilgrastim: a pilot study in Europe. Journal of Drug Assessment, 8(1), pp.70-76.

[15] Ball, R., Toh, S., Nolan, J., Haynes, K., Forshee, R. and Botsis, T., 2018. Evaluating automated approaches to anaphylaxis case classification using unstructured data from the FDA Sentinel System. Pharmacoepidemiology and drug safety, 27(10), pp.1077-1084.

[16] Goyal, N., Tsivgoulis, G., Malhotra, K., Katsanos, A.H., Pandhi, A., Alsherbini, K.A., Chang, J.J., Hoit, D., Alexandrov, A.V., Elijovich, L. and Fiorella, D., 2019. Minimally invasive endoscopic hematoma evacuation vs best medical management for spontaneous basal-ganglia intracerebral hemorrhage. Journal of NeuroInterventional Surgery, 11(6), pp.579-583.

[17] Herzberg, S., Hansen, M., Schoonover, A., Skarica, B., McNulty, J., Harrod, T., Snowden, J.M., Lambert, W. and Guise, J.M., 2019. Association between measured teamwork and medical errors: an observational study of prehospital care in the USA. BMJ open, 9(10), p.e025314.

[18] Webb, J., Shah, L.D. and Lynch, H.F., 2020. Ethically allocating COVID-19 drugs via pre-approval access and emergency use authorization. The American Journal of Bioethics, 20(9), pp.4-17.

[19] Wiese, A.D., Griffin, M.R., Stein, C.M., Schaffner, W., Greevy, R.A., Mitchel Jr, E.F. and Grijalva, C.G., 2018. Validation of discharge diagnosis codes to identify serious infections among middle age and older adults. BMJ open, 8(6), p.e020857.

[20] Jorstad, K.T., 2020. Intersection of artificial intelligence and medicine: tort liability in the technological age. Journal of Medical Artificial Intelligence, 3.

[21] Eloranta, S. and Boman, M., 2022. Predictive models for clinical decision making: Deep dives in practical machine learning. *Journal of internal medicine*, *292*(2), pp.278-295.

[22] Szalados, J.E., 2021. Statutory Controls and Regulation of Pharmaceuticals and Medical Devices. In The Medical-Legal Aspects of Acute Care Medicine: A Resource for Clinicians, Administrators, and Risk Managers (pp. 547-571). Cham: Springer International Publishing.

[23] Mouchantaf, R., Auth, D., Moride, Y., Raine, J., Han, S.Y. and Smith, M.Y., 2021. Risk management for the 21st century: current status and future needs. Drug Safety, 44(4), p.409.

[24] Holbrook, A.M., Vidug, K., Yoo, L., Troyan, S., Schulman, S., Douketis, J., Thabane, L., Giilck, S., Koubaesh, Y., Hyland, S. and Keshavjee, K., 2022. Coordination of Oral Anticoagulant Care at Hospital Discharge (COACHeD): protocol for a pilot randomised controlled trial. Pilot and Feasibility Studies, 8(1), p.166.

[25] Bajandas, F.F. and Ray, G.K., 2018. Implementation and Use of Electronic Case

Management Systems in Federal Agency Adjudication. In Administrative Conference Recommendation (Vol. 3, pp. 44-49).

[26] Bommasani, R., Hudson, D.A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M.S., Bohg, J., Bosselut, A., Brunskill, E. and Brynjolfsson, E., 2021. On the opportunities and risks of foundation models. arXiv preprint arXiv:2108.07258.

[27] Sharma, A., Mahaffey, K.W., Gibson, C.M., Hicks, K.A., Alexander, K.P., Ali, M., Chaitman, B.R., Held, C., Hlatky, M., Jones, W.S. and Mehran, R., 2022. Clinical events classification (CEC) in clinical trials: report on the current landscape and future directions—proceedings from the CEC Summit 2018. American Heart Journal, 246, pp.93-104.

[28] Coravos, A., Goldsack, J.C., Karlin, D.R., Nebeker, C., Perakslis, E., Zimmerman, N. and Erb, M.K., 2019. Digital medicine: a primer on measurement. Digital biomarkers, 3(2), pp.31-71.

[29] Coravos, A., Goldsack, J.C., Karlin, D.R., Nebeker, C., Perakslis, E., Zimmerman, N. and Erb, M.K., 2019. Digital medicine: a primer on measurement. Digital biomarkers, 3(2), pp.31-71.

[30] Herbst, R.B., Margolis, K.L., McClellan, B.B., Herndon, J.L., Millar, A.M. and Talmi, A., 2018. Sustaining integrated behavioral health practice without sacrificing the continuum of care. Clinical Practice in Pediatric Psychology, 6(2), pp.117-128.