International Journal of Electrical, Electronics and Computer Systems
ISSN: 2347-2820
Volume 15 Issue 1s (2026) | Pages: 71-76

AI-Based Intelligent Insurance Agent Using Retrieval-Augmented Generation

Sakshi Alamkhane[1], Aditi Anarse[2], Mahek Shaikh[3], Nandini Chavan[4], Rahul Korke[5]

[1,2,3,4,5] Department of Computer Engineering, Genba Sopanrao Moze College Of Engineering, Balewadi, Pune-45

Peer Review Information

Type: *Article*
Received: *13 February 2026*
Revised: *14 March 2026*
Accepted: *15 April 2026*
Published: *19 May 2026*

Abstract

The insurance industry faces significant challenges in providing efficient, accurate, and personalized customer support due to the complexity of policies and reliance on human agents. This research presents an AI-Based Intelligent Insurance Agent that utilizes Retrieval-Augmented Generation (RAG) to deliver context-aware and reliable responses. The system integrates natural language processing, machine learning, and database-driven retrieval to assist users in understanding insurance policies, recommending suitable plans, and supporting claim-related queries. It offers both text and voice interaction, ensuring 24/7 accessibility and improved user engagement. The methodology involves retrieving relevant data from a structured database and generating responses using an AI model. The proposed system reduces operational costs, enhances accuracy, and improves customer satisfaction. The results demonstrate that the system provides faster and more consistent support compared to traditional methods. This research highlights the potential of AI in transforming the insurance sector through automation and intelligent decision-making.

Keywords: AI Agent; Insurance Technology; RAG; Chatbot; NLP

How to Cite This Article

Alamkhane, S., Anarse, A., Shaikh, M., Chavan, N., & Korke, R. (2026). *AI-Based Intelligent Insurance Agent Using Retrieval-Augmented Generation*. *International Journal of Electrical, Electronics and Computer Systems,* 15(1s), 71–76.

Introduction

The insurance sector is undergoing rapid digital transformation, yet many processes remain inefficient due to reliance on manual operations. Customers often find insurance policies difficult to understand, and traditional customer service systems are slow and costly. Existing chatbots provide only limited support due to their lack of contextual understanding.

Artificial Intelligence (AI) has emerged as a powerful tool to address these challenges by enabling automated, intelligent, and scalable solutions. The proposed AI-Based Insurance Agent leverages conversational AI and Retrieval-Augmented Generation to provide accurate, real-time assistance to users. The system enhances transparency, reduces dependency on human agents, and improves customer experience.

The main objective of this study is to develop a smart insurance assistant capable of understanding user queries, recommending policies, and assisting with claims. This research contributes to improving efficiency, reducing costs, and enhancing accessibility in the insurance domain.

Literature Review

Bhattacharya et al. (2025) studied the role of Artificial Intelligence in the insurance sector, focusing on underwriting, claims processing, and risk assessment. The study concluded that AI improves operational efficiency and decision-making but faces challenges related to data quality, ethical concerns, and regulatory compliance.

Ercole (2025) proposed a Retrieval-Augmented Generation (RAG) based AI system for insurance knowledge management. The system demonstrated improved contextual understanding and response accuracy compared to traditional rule-based chatbots.

Bradley et al. (2025) developed an LLM-based insurance assistant system that helps users retrieve policy-related information. The model achieved high accuracy in information retrieval and improved customer interaction using natural language processing techniques.

Hanmante et al. (2025) designed a multi-module AI system for health insurance using RAG architecture. The system integrates policy recommendation, query handling, and conversational assistance, improving user decision-making and interaction efficiency.

Sawarkar et al. (2024) explored hybrid RAG models combining semantic search and vector-based indexing for insurance applications. The approach improved handling of complex queries but required higher computational resources.

Hillebrand et al. (2024) introduced a RAG-based chatbot for regulatory compliance in financial and insurance systems. The system enhanced response reliability and contextual accuracy but depended heavily on the quality of retrieved data.

Kuppan et al. (2024) presented a survey on AI applications in insurance and real estate, highlighting its use in fraud detection, risk assessment, and customer service automation. The study also identified challenges such as explainability and data bias.

The Geneva Association Report (2025) analyzed the impact of generative AI on the insurance customer journey. It concluded that AI improves personalization, customer engagement, and operational efficiency in insurance services.

Recent industry research (2026) showed that AI-powered virtual agents significantly enhance customer support by reducing response time and operational costs while improving overall customer satisfaction.

A recent study on RAG-enhanced chatbot architectures (2025) demonstrated that integrating retrieval mechanisms with generative AI improves accuracy, reduces hallucination, and enhances scalability in insurance systems, though challenges like data privacy remain.

Methodology

Describe research design, data collection methods, tools used, and sample size. This research adopts a system design and experimental approach to develop an AI-Based Insurance Agent using Retrieval-Augmented Generation (RAG). The methodology focuses on building an intelligent conversational system capable of understanding user queries and providing accurate, context-aware responses.

Research Design

The proposed system follows a hybrid AI model combining Natural Language Processing (NLP) and information retrieval techniques. The

architecture is based on Retrieval-Augmented Generation, where relevant insurance data is retrieved from a database and then processed using a generative AI model to produce meaningful responses. The system is designed to support both text and voice-based interactions, ensuring accessibility and usability.

Data Collection Methods

Data for the system is collected from multiple reliable sources, including:

- Publicly available insurance datasets
- Policy documents from insurance providers
- Sample customer queries and FAQs

The collected data undergoes preprocessing steps such as:

- Removal of null and duplicate values
- Text normalization (tokenization and lemmatization)
- Feature extraction (policy type, user intent, keywords) The processed dataset is then divided into:
- Training Data (80%)
- Testing Data (20%)

Tools and Technologies Used

The system is implemented using the following tools and technologies:

- Programming Language: Python
- Frontend: React.js
- Backend: Flask
- Database: MongoDB
- AI Techniques: Natural Language Processing (NLP), Machine Learning
- Model Approach: Retrieval-Augmented Generation (RAG)

System Workflow

1. User enters a query through text or voice input.
2. The system processes the query using NLP techniques.
3. Relevant information is retrieved from the database.
4. The AI model generates a context-aware response using retrieved data.
5. The response is delivered to the user in real-time.

Sample Size

The system is trained and tested on a dataset consisting of:

- Insurance policy records
- Customer interaction samples
- Frequently asked questions (FAQs)

The dataset includes hundreds to thousands of records, ensuring sufficient data for training and evaluation of the model.

Evaluation Metrics

The performance of the system is evaluated based on:

- Accuracy of responses
- Response time
- User satisfaction

- Relevance of recommendations

Results / Findings

The proposed AI-Based Insurance Agent was evaluated based on response accuracy, query handling efficiency, recommendation relevance, and system performance. The results demonstrate that the system provides accurate, fast, and context-aware responses compared to traditional methods.

Performance Evaluation

Table 1. Performance Evaluation Metrics of AI-Based Intelligent Insurance Agent System

Metric	**Result (%)**	**Description**
Response Accuracy	91%	Correctness of answers generated by the system
Query Resolution Rate	88%	Queries successfully handled without human intervention
Recommendation Accuracy	85%	Relevance of insurance policies suggested
User Satisfaction Score	90%	Based on user feedback and testing
Response Time	< 2 sec	Average time to generate response

Comparison with Traditional Systems

Table 2. Comparison Between Traditional Insurance Systems and Proposed AI-Based Insurance System

Parameter	**Traditional System**	**Proposed AI System**
Availability	Limited (Working Hours)	24/7 Available
Response Speed	Slow	Fast (<2 sec)
Accuracy	Moderate	High
Personalization	Low	High
Operational Cost	High	Reduced

Query Handling Analysis

Table 3. Query Handling Performance of the AI-Based Insurance Agent System

Query Type	**Total Queries**	**Successfully Handled**	**Success Rate**
Policy Information	120	112	93%
Claim Assistance	100	87	87%
General Queries	80	72	90%

Key Findings

- The system achieved high response accuracy (above 90%) using RAG.
- Query resolution improved significantly compared to manual systems.
- Personalized recommendations increased user satisfaction.

- Response time was consistently under 2 seconds, ensuring real-time interaction.
- The system reduced dependency on human agents and operational workload.

Observations

- Retrieval-Augmented Generation improved contextual understanding.
- Performance depends on the quality of the dataset.
- The system performs best for structured insurance-related que

Discussion

The results of the proposed AI-Based Insurance Agent indicate that the integration of Retrieval-Augmented Generation (RAG) with conversational AI significantly enhances system performance. The high response accuracy and fast response time demonstrate that the system effectively fulfills the research objective of providing real-time and context-aware assistance. The system was able to handle a majority of user queries without human intervention, confirming its capability to reduce dependency on manual customer support while maintaining service quality.

The findings also show that the recommendation engine provides relevant and personalized insurance suggestions, which improves user satisfaction and decision-making. These results are consistent with previous studies such as Hillebrand et al. [6] and Bradley et al. [3], which emphasize the effectiveness of RAG-based systems in improving chatbot accuracy. Additionally, the system outperforms traditional insurance support systems in terms of availability, speed, and scalability, supporting the need for AI-driven digital transformation in the insurance sector.

However, certain limitations were observed during the study. The performance of the system is highly dependent on the quality and availability of data, and inaccuracies may occur when handling incomplete or ambiguous queries. This limitation aligns with prior research [10], which highlights challenges related to data dependency and system integration. Overall, the proposed system successfully meets the research objectives, but further improvements in data quality, scalability, and explainability are required for real-world implementation.

Conclusion

The proposed AI-Based Insurance Agent successfully demonstrates the potential of integrating Artificial Intelligence with Retrieval-Augmented Generation (RAG) to enhance customer support in the insurance sector. The system provides accurate, fast, and context-aware responses to user queries while offering personalized policy recommendations. The results indicate that the solution improves response accuracy, reduces query resolution time, and ensures 24/7 availability, thereby addressing the limitations of traditional manual systems.

The key takeaway from this research is that AI-driven conversational agents can significantly reduce operational costs and dependency on human agents while improving customer satisfaction and service efficiency. The use of RAG enables better contextual understanding compared to traditional chatbots, making the system more reliable and scalable for real-world applications.

However, the system has certain limitations, including dependency on high-quality datasets and computational resources. In cases of incomplete or ambiguous data, the accuracy of responses may be affected. Future research can focus on integrating real-time insurance APIs, enhancing multilingual support, and improving explainability of AI models. Additionally, incorporating advanced features such as fraud detection, voice-based interaction, and mobile application deployment can further enhance the system's capabilities and practical usability.

References

1. Bhattacharya, S., et al., "AI Revolution in Insurance: Bridging Research and Reality," *Frontiers in Artificial Intelligence*, 2025.
2. Hanmante, S., Patil, S., & Shahade, A., "A Multi-Module AI System for Intelligent Health Insurance Support Using RAG," *Scientific Reports*, 2025.
3. Bradley, A., Hastings, J., & Ahmed, K., "Axlerod: An LLM-Based Chatbot for Assisting Insurance Agents," *arXiv*, 2025.
4. Huang, D. H., et al., "RAG-Enhanced AI Agent Chatbot Architecture for Optimizing Insurance Operations," 2025.
5. Son, N., et al., "Development and Evaluation of a Retrieval-Augmented Generation Chatbot System," 2025.
6. Neha, F., et al., "Retrieval-Augmented Generation in Healthcare: A Systematic Review," *MDPI*, 2025.
7. "Privacy Challenges and Solutions in RAG-Enhanced LLMs," *arXiv*, 2025.

8. “Application Paper on the Supervision of Artificial Intelligence in Insurance,” *IAIS Report*, 2025.
9. “Gen AI in the Insurance Customer Journey,” *The Geneva Association*, 2025.
10. “AI-Powered Life Insurance Claims Adjudication Using LLMs and RAG,” *IJSRA*, 2025.