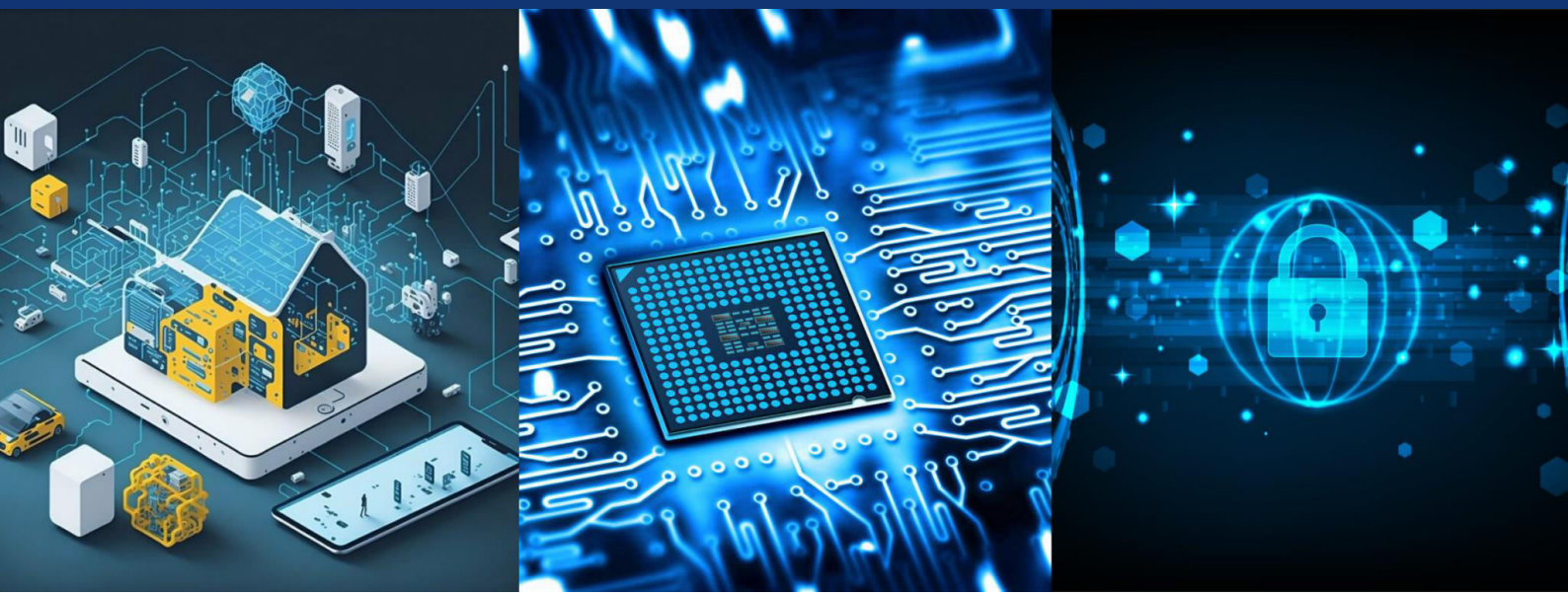
IJIRCCE

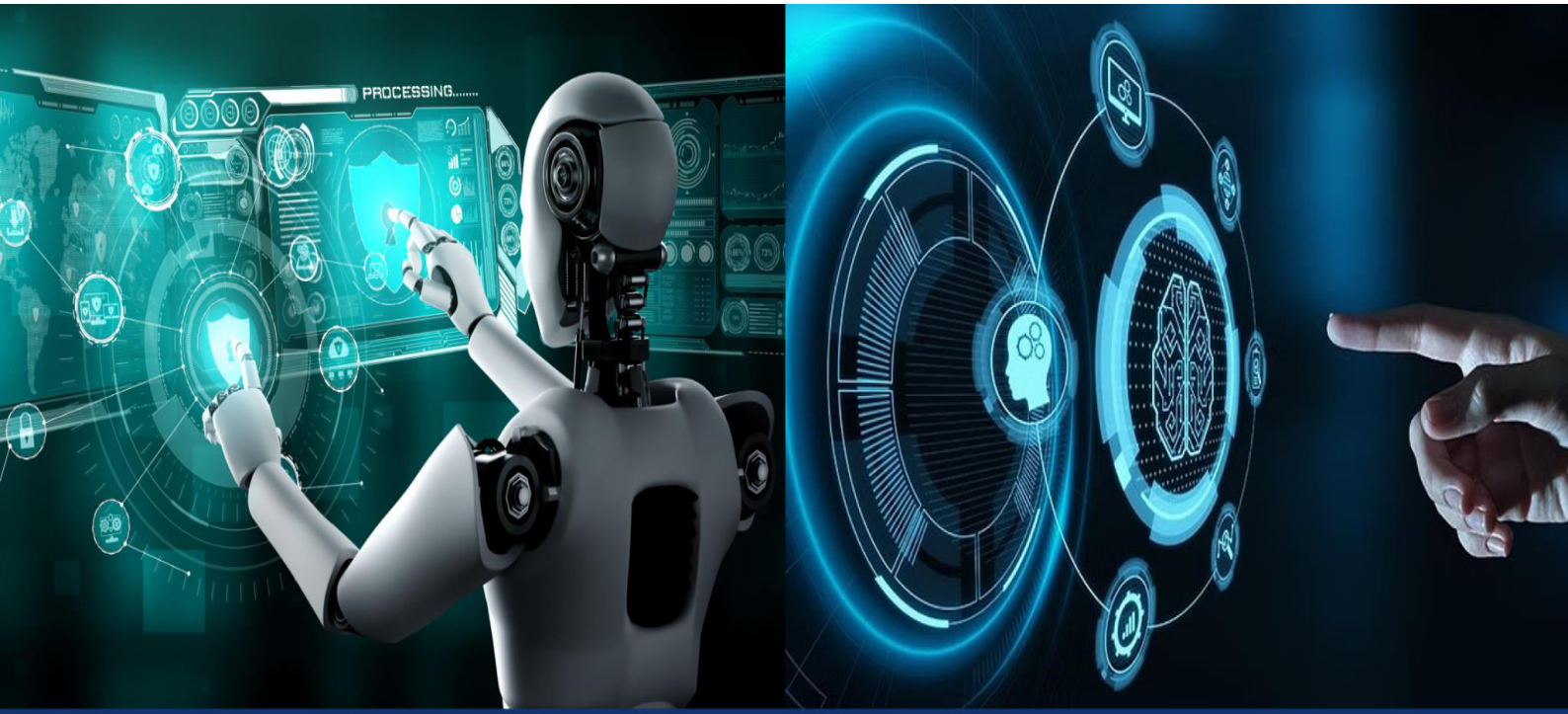
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

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



Impact Factor: 8.625

Volume 13, Issue 1, January 2025

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301001

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

Revolutionizing IVR Systems with Generative AI for Smarter Customer Interactions

Vigneshwaran Jagadeesan Pugazhenthi, Jarvish Kisanth Singh, Gokul Pandy

IEEE Member, Henrico, Virginia, USA

Researcher, Tamil Nadu, India

IEEE Senior Member, Henrico, Virginia, USA

ABSTRACT: The incorporation of Generative AI into Interactive Voice Response (IVR) systems is revolutionizing the way businesses engage with customers, significantly boosting both operational efficiency and customer satisfaction. [1] Unlike traditional IVR systems, which rely on rigid, pre-recorded responses, modern AI-driven solutions offer more dynamic, adaptable, and intuitive interactions. These systems can process and generate natural language, enabling highly personalized, context-aware conversations that are more fluid and less frustrating for users. This paper examines the transformative potential of Generative AI in enhancing IVR systems, focusing on advancements in speech recognition, real-time customer need adaptation, and reduction of wait times. It also explores the technological, ethical, and implementation challenges faced by organizations while showcasing successful case studies of businesses integrating Generative AI into their IVR platforms. The outcome is a streamlined, scalable, and user-centric experience that meets the evolving demands of today's consumers..

KEYWORDS: Generative AI, IVR , NLU, Artificial Intelligence, Customer Experience

I. INTRODUCTION

Interactive Voice Response (IVR) systems have been a key part of customer service, helping businesses handle large numbers of customer interactions. Traditionally, these systems rely on fixed menus and pre-recorded responses, which can often feel frustrating and impersonal. As customers now expect more personalized and natural interactions, there is a growing need for smarter solutions that can meet these demands.[2]

Generative Artificial Intelligence (AI) presents a promising solution. Unlike traditional IVR systems that follow set paths, Generative AI can understand natural language, learn from each interaction, and create responses in real-time. This enables more intuitive, personalized experiences that can adjust to each customer's needs, shifting IVR systems from rigid, scripted exchanges to more flexible, conversational interactions tailored to individual preferences.

This paper explores how Generative AI is reshaping IVR systems, focusing on its effects on customer satisfaction, operational efficiency, and technological progress. It examines the advancements and challenges of implementing this technology and discusses the ethical implications of AI integration. Through case studies, we will showcase organizations that have successfully integrated Generative AI to transform their IVR systems, leading to a more scalable and user-friendly customer service experience.

II. RELATED WORK

The integration of Generative AI in Interactive Voice Response (IVR) systems is a relatively new field, but there has been a growing body of research and practical implementations that demonstrate its potential. Several studies and initiatives have explored how AI and machine learning can enhance traditional IVR systems, improving both user experience and operational efficiency. Speech recognition and NLP technologies[3] have been central to the evolution of IVR systems. Traditional IVR systems were limited by keyword-based recognition, but AI-based systems use more advanced techniques like deep learning and natural language processing to understand and interpret customer queries more effectively. Research by **Hinton et al. (2012)** on deep neural networks for speech recognition has paved the way for these advancements, making AI-driven IVR systems more accurate and responsive. The introduction of Generative

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

AI in voice-based interactions has garnered attention for its potential to create more human-like conversations. **Vaswani et al. (2017)** introduced the Transformer model, a significant breakthrough in generative models, which has since been adapted for voice synthesis and recognition tasks. Studies such as **Li et al. (2020)** show that Generative AI can not only recognize speech but also generate contextually appropriate responses, which is crucial for IVR systems to move from static to dynamic, context-aware interaction

III. PROPOSED DESIGN

A. Design Considerations:

- IVR system gets the speech input from the customer
- System converts the input using Speech to Text technology
- System identifies the intent using NLU/NLP technologies
- The intent is then fed into Generative AI Model like Jamba 1.5 Mini Model within Amazon Bedrock
- The response from model is then converted into life like voice form using services like Amazon Polly[10]
- The response from model is then spoken to the customer

B. Description of the Proposed Design:

Step 1: IVR system gets the speech input from the customer :
The IVR system receives speech input from the customer, capturing their spoken words through a microphone or phone [8]. This audio input is then processed for further conversion into text for analysis.

Step 2: System converts the input using Speech to Text technology :
The system uses **Speech-to-Text (STT)** technology to convert the customer's spoken words into written text. This process involves recognizing the sounds in the speech and mapping them to corresponding text. Once converted, the text is ready for analysis, enabling the system to understand the customer's request, intent, and relevant details for further processing. This step ensures that the system can handle a wide range of speech patterns and languages accurately

Step 3: System identifies the intent using NLU/NLP technologies
The system identifies the customer's intent using **Natural Language Understanding (NLU)** and **Natural Language Processing (NLP)** technologies. These technologies analyze the text input to determine the meaning behind the words, recognizing key phrases, context, and the specific request or action the customer wants. By processing the language, the system can categorize the intent and extract essential information, such as customer inquiries or service requests, to generate an appropriate response .

Step 4 : The intent is then fed into Generative AI Model like Jamba 1.5 Mini Model within Amazon Bedrock
The identified intent is then fed into a **Generative AI Model**,[4] such as **Jamba 1.5 Mini Model** within **Amazon Bedrock**. This model leverages advanced machine learning algorithms to generate a natural and contextually relevant response based on the customer's intent. By understanding the nuances of the input, the model crafts personalized replies, enabling the IVR system to engage in more dynamic, conversational exchanges that meet the customer's needs.

Step 5 : The response from model is then converted into life like voice form using services like Amazon Polly
The response generated by the model is then converted into lifelike speech using services like **Amazon Polly**. Amazon Polly uses advanced **Text-to-Speech (TTS)** technology[5] to synthesize the generated text into natural-sounding voice output, with various voice options and language capabilities. This ensures the IVR system delivers a human-like, engaging, and clear response, providing an enhanced customer experience.

Step 6 : The response from model is then spoken to the customer
The response generated by the model is then spoken to the customer through the IVR system. Using **Text-to-Speech (TTS)** technology [5], the system converts the text into audible speech, delivering the response in a clear and natural-sounding voice to the customer, ensuring effective communication and a seamless experience

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301001

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

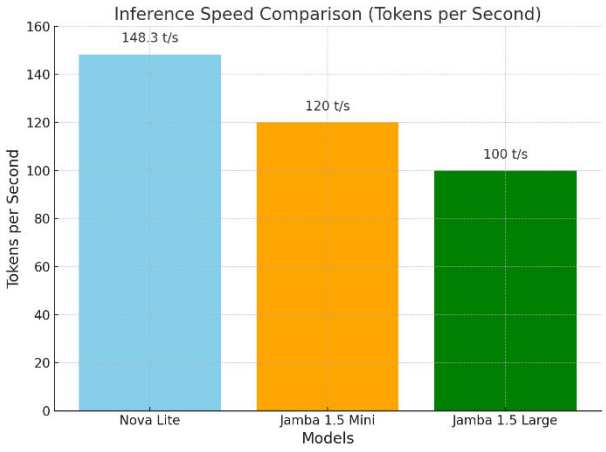
IV. EXPERIMENT OUTCOMES

Below are some of the experimental outcomes from different models with AWS Bedrock [6] for the same input from the customers. You can see how the model responds differently for each input. For our experiment we are considering Nova Lite, Jamba 1.5 Mini and Jamba 1.5 Large. Novalite is an Amazon model. Jamba models are AI21 Labs

Model Name	Input by the User	Response
Nova Lite	Tell me about health insurance claims	Health Insurance Claims are formal requests made by a policy
Jamba 1.5 Mini V1	Tell me about health insurance claims	Health insurance claims are formal requests made by an insurance
Jamba 1.5 Large v1	Tell me about health insurance claims	Health insurance claims are requests made to health insurance providers

The inference speed of these models vary

Nova Lite: Approximately 148.3 tokens per second.[9]
Jamba 1.5 Mini: Approximately 120 tokens per second.
Jamba 1.5 Large : Approximately 100 tokens per second. Slower than mini but handles complex tasks



Feature	Nova Lite	Jamba 1.5 Mini	Jamba 1.5 Large
Inference Speed	Optimized for speed, but not as high as Jamba Mini	Over 150 tokens per second for 10,000 tokens	Slower than Mini but optimized for complex tasks
Context Window	Smaller context window, typically a few thousand tokens	256K tokens	256K tokens
Model Size	Lightweight, optimized for real-time applications	Smaller model, optimized for speed	Larger model, optimized for complex reasoning
Best Use Case	Lightweight tasks, mobile/edge applications requiring moderate context	High-demand tasks with long inputs (e.g., document summarization, real-time Q&A)	Complex reasoning, tasks requiring deep context and large inputs
Output Quality	Good quality, but may struggle with complex	Good for most tasks, may not handle deep reasoning	Excellent quality for complex tasks requiring

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

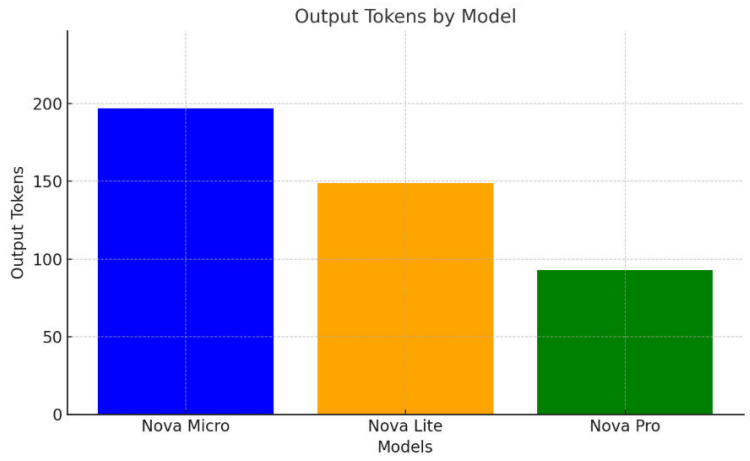
International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

	reasoning compared to larger models	as well as larger models	deep reasoning
Optimization	Optimized for efficiency, ideal for smaller tasks	Optimized for fast, high-throughput tasks	Optimized for quality and reasoning with larger prompts
Use in Applications	Suitable for real-time, smaller-scale tasks with moderate context requirements	Ideal for tasks needing large context windows and fast responses	Best for complex applications needing higher-quality and deeper understanding

Output speed of Amazon Models :

An Output Speed defines Tokens per second received while the model is generating token[7]



V. CONCLUSION AND FUTURE WORK

The future of generative AI-powered IVR systems lies in expanding their capabilities to handle more complex and dynamic customer scenarios while integrating seamlessly with broader business ecosystems. Key areas of focus include enhancing emotional intelligence through sentiment analysis, enabling these systems to detect customer emotions and respond empathetically. Additionally, the development of multimodal interaction capabilities—such as integrating voice, text, and visual elements—can provide a more comprehensive and unified customer experience. Advancing localization features to support diverse languages and ensuring inclusivity for customers with disabilities will be crucial for reaching a global audience. Furthermore, prioritizing robust data privacy and security measures will help safeguard sensitive customer information and ensure compliance with evolving regulations. Finally, the integration of predictive analytics could enable IVR systems to anticipate customer needs and deliver proactive support, creating a more personalized and efficient service experience. These advancements promise to redefine customer interactions, driving innovation and setting new standards for customer service excellence.

REFERENCES

1. K. S. Kaswan, J. S. Dhatterwal, K. Malik and A. Baliyan, "Generative AI: A Review on Models and Applications," 2023 International Conference on Communication, Security and Artificial Intelligence (ICCSAI), Greater Noida, India, 2023, pp. 699-704, doi: 10.1109/ICCSAI59793.2023.10421601.
2. K. S. Kaswan, J. S. Dhatterwal, K. Malik and A. Baliyan, "Generative AI: A Review on Models and Applications," 2023 International Conference on Communication, Security and Artificial Intelligence (ICCSAI), Greater Noida, India, 2023, pp. 699-704, doi: 10.1109/ICCSAI59793.2023.10421601.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

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

3. A. Gelbukh, "Natural language processing," Fifth International Conference on Hybrid Intelligent Systems (HIS'05), Rio de Janeiro, Brazil, 2005, pp. 1 pp.-, doi: 10.1109/ICHIS.2005.79.
4. M. Jovanović and M. Campbell, "Generative Artificial Intelligence: Trends and Prospects," in Computer, vol. 55, no. 10, pp. 107-112, Oct. 2022, doi: 10.1109/MC.2022.3192720.
5. V. M. Reddy, T. Vaishnavi and K. P. Kumar, "Speech-to-Text and Text-to-Speech Recognition Using Deep Learning," 2023 2nd International Conference on Edge Computing and Applications (ICECAA), Namakkal, India, 2023, pp. 657-666, doi: 10.1109/ICECAA58104.2023.10212222.
6. Shikhar Kwatra; Bunny Kaushik, Generative AI with Amazon Bedrock: Build, scale, and secure generative AI applications using Amazon Bedrock , Packt Publishing, 2024.
7. J. Conde et al., "Speed and Conversational Large Language Models: Not All Is About Tokens per Second" in Computer, vol. 57, no. 08, pp. 74-80, Aug. 2024, doi: 10.1109/MC.2024.3399384.
8. Medallia. "Interactive Voice Response (IVR)." Medallia, www.medallia.com/experience-101/glossary/interactive-voice-response/. Accessed 6 Jan. 2025
9. Artificial Analysis. (n.d.). Nova Lite Providers. Retrieved January 6, 2025, from https://artificialanalysis.ai/models/nova-lite/providers
10. A. Werchniak et al., "Exploring the application of synthetic audio in training keyword spotters," ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 2021, pp. 7993-7996, doi: 10.1109/ICASSP39728.2021.9413448

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details