AI-Based Voice Agent for Automated Sales Calls

Ahmed Amarak
School of Science and Engineering
Al Akhawayn University in Ifrane
Ifrane, Morocco
A.Amarak@aui.ma

Mohamed Igamane
School of Computer Science
Nanjing University of Posts
and Telecommunications
Nanjing, China
F21030210@njupt.edu.cn

Tajjeeddine Rachidi
School of Science and Engineering
Al Akhawayn University in Ifrane
Ifrane, Morocco
T.Rachidi@aui.ma

Yousra Chtouki
School of Science and Engineering
Al Akhawayn University in Ifrane
Ifrane, Morocco
Y.Chtouki@aui.ma

Abstract — **This paper presents an AI-based voice agent designed to enhance customer engagement in automated sales. Operating continuously, the agent leverages the power of Large Language Models (LLMs) to interact with potential customers, persuading them to purchase products or schedule appointments with human representatives. Initial performance evaluation gives extremely positive results highlighting the agent's potential to revolutionize sales interactions and improve operational efficiency. These are demonstrated by a conversion rate of 8%, significantly higher than industry averages of 2-5%, high ratings for friendliness, clarity, and realism in customer satisfaction surveys, and an average of 1.7 seconds reaction time during conversation.**

Keywords— AI-based call agent, sales automation, customer engagement, NLP, LLM.

I. INTRODUCTION

In today's highly competitive market, businesses face a number of difficult challenges that make customer engagement more important than ever [1]. Companies across various industries rely on customer and sales agents to reach potential clients through outbound calls, to promote products or services, and to drive revenue growth. These agents play a pivotal role in personalizing the customer experience and building lasting relationships. However, employing human sales agents comes with significant challenges and costs.

Human sales agents are not only expensive to hire and train, but they also have inherent limitations. They can typically work only 8 hours a day, which restricts the number of potential customers they can reach within a given time frame. Additionally, the quality and consistency of customer interactions can vary significantly between agents, affecting the overall effectiveness of the sales process. Research indicates that AI-based conversational agents can enhance operational performance and reduce costs in call centers by providing continuous and consistent customer service [2], [3].

To address these challenges, we propose an AI-based call agent designed to automate customer engagement and sales processes. This AI agent leverages advanced natural language processing (NLP) techniques to interact with potential customers, providing a scalable and cost-effective solution. Unlike human agents, the AI-based call agent can operate 24/7, reaching a much larger audience without the constraints of working hours. It ensures consistent and personalized interactions with customers, enhancing their experience and increasing the likelihood of successful engagements. Studies have shown that AI-based voice assistants can significantly influence consumer behavior and enhance customer service efficiency [4]. Additionally, AI voice bots are found to be effective in providing personalized recommendations and assisting in decision-making processes, which further boosts customer satisfaction and engagement [5].

By addressing the limitations of human sales agents and leveraging the capabilities of AI, our proposed solution aims to revolutionize the way companies engage with potential customers, ultimately enhancing efficiency and driving sales growth.

This paper is organized as follows: In Section 2, we review related work in the field of AI-driven sales automation and customer engagement tools. Section 3 presents the system design and architecture of the AI-based call agent, detailing the system components and the call structure. Section 4 covers the implementation aspects, including the technologies used and an evaluation of the system's performance, such as experimental setup, metrics, and results. Section 5 discusses the limitations and future work, highlighting areas for improvement and potential advancements. Finally, Section 6 concludes the paper with a summary of findings and broader implications.

II. RELATED WORK

In this section, we will review existing solutions, highlight comparisons with other voice-based sales agents, and identify gaps in current technologies, emphasizing how our proposed solution addresses these gaps.

These solutions range from chatbots that handle customer inquiries to sophisticated virtual sales agents that manage entire sales funnels. For instance, voice-activated shopping devices like Amazon's Alexa and Alibaba's Tmall Genie have gained popularity, offering a new channel for consumers to make purchases through voice commands [6].

Studies have shown that the adoption of voice AI significantly affects consumer behavior, leading to increased spending and enhanced shopping experiences. Research on Alibaba's Tmall Genie revealed that consumers who adopted the voice AI device spent 23% more on average, translating to a substantial increase in sales revenue [7]. The impact of voice AI is particularly pronounced for products that do not require active search or comparison, such as items with low substitutability or high purchase frequency [7].

Voice-based sales agents offer unique advantages over traditional chatbots and text-based systems. They provide a more natural and intuitive way for customers to interact with technology, replicating the experience of speaking with a human sales agent. Voice AI systems can handle complex

queries, provide personalized recommendations, and assist with purchasing decisions in real-time.

The integration of AI in voice-based assistants for digital marketing has further expanded the capabilities of these tools. These assistants can engage users in meaningful dialogues, answer questions, and even guide users through complex decision-making processes [8]. The effectiveness of voice-based assistants in driving sales and enhancing customer engagement has been documented in various studies, highlighting their potential to revolutionize the e-commerce landscape [9].

Beyond sales, AI technologies have been successfully applied to other business processes, such as human resources management. [10] examined the role of AI in optimizing recruitment and selection processes, demonstrating how automation and expert systems improve operational efficiency and minimize human bias. Their findings underscore AI's capacity to take over repetitive tasks, such as candidate sourcing and screening, ensuring consistent and high-quality outcomes. These insights align with the principles underpinning AI-based call agents, which similarly aim to automate routine interactions while maintaining a high standard of personalization and engagement.

Despite the advancements in AI-driven sales automation, there are still notable gaps in the current technologies. Many existing solutions lack the sophistication needed to fully replicate the persuasive and negotiating capabilities of human sales agents. Additionally, while chatbots and text-based systems have been widely adopted, there is a need for more advanced voice-based solutions that can operate 24/7, provide consistent interactions, and handle a higher volume of customer inquiries.

Our proposed AI-based call agent addresses these gaps by using state-of-the-art NLP techniques to engage customers in meaningful dialogues. Unlike traditional systems, our voice sales agent can operate continuously, ensuring that businesses are always reaching out to potential customers. Our solution offers a scalable and cost-effective alternative to human sales agents, capable of handling complex interactions and providing personalized recommendations.

III. SYSTEM DESIGN AND ARCHITECTURE

In this section, we will delve into the system architecture and the call structure of our AI-based call agent. We will outline each system component's role and interaction with others to provide a comprehensive understanding of the system's functionality. Following this, the AI agent call structure will be detailed, explaining how the AI guides potential customers through a well-organized sales conversation to ensure effective engagement and persuasive interactions.

A. System Architecture

The AI-based call agent system consists of several key components that work together to facilitate seamless and efficient customer interactions (see Fig. 1). The main components include a telephony API, a speech-to-text

model, a text-to-speech model, and an LLM. The interaction between these components is outlined below.

The telephony API acts as the interface between the AI system and the customer. It is responsible for initiating outbound calls to potential customers, converting audio signals to a format that can be processed by the speech-to-text model. The telephony API ensures that the audio data from the customer's end is accurately transmitted to the system for further processing, enabling the AI to engage with customers effectively.

The speech-to-text model is responsible for converting the audio input from the customer into text. This model tries to accurately transcribe spoken language into written form. The text output is then passed to the LLM for further processing.

The LLM is the core component that processes the text input and generates appropriate responses. It leverages machine learning algorithms to understand the context of the conversation and formulate relevant and persuasive replies. The LLM's capabilities enable it to handle a wide range of customer inquiries, from simple questions to complex queries about product integration.

Once the LLM generates a textual response, the text-to-speech model converts it back into audio. This model ensures that the synthesized speech sounds natural and engaging, enhancing the overall customer experience. The audio output is then transmitted back to the customer through the telephony API.
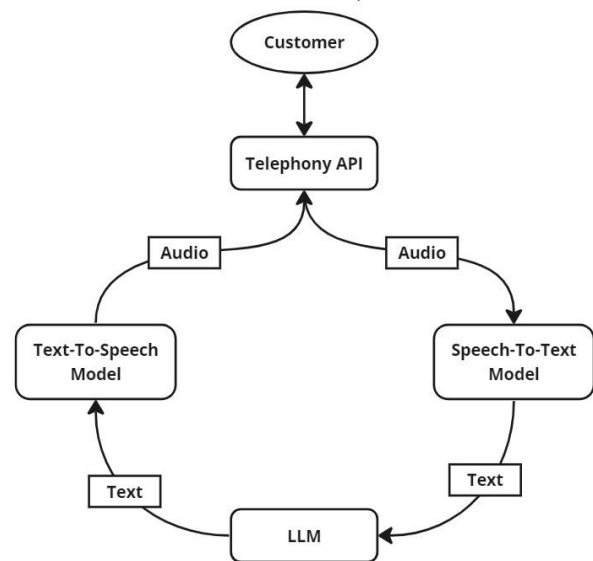

Fig. 1. System Architecture of the AI-based call agent

B. Call Structure

The following call structure is designed to systematically guide potential customers through an enriching sales conversation, ensuring a comprehensive and persuasive interaction (see Fig. 2).

The process begins with an initial greeting, where the AI agent builds rapport with the potential customer and introduces the purpose of the call. By using a friendly and

casual greeting, the agent briefly mentions the product or service, setting a positive tone for the conversation.

Following the initial greeting, the AI agent conducts a needs discovery phase. During this phase, the agent asks open-ended questions to understand the customer's needs and interests, identifying any pain points or goals the customer may have. Based on these identified needs, the AI agent presents the product or service, highlighting its benefits. This involves providing a brief overview of the key features and advantages tailored to the customer's specific requirements. If the customer has any questions or concerns, the AI agent addresses them by listening attentively and providing clear, concise answers, which helps build trust and credibility.

After addressing any concerns, the AI agent explains the unique value proposition of the product or service, emphasizing what sets it apart from competitors and how it can solve the customer's problems. With a clear understanding of the customer's needs, the AI agent offers a tailored solution, suggesting specific ways the product or service can benefit the customer. Finally, the AI agent encourages the customer to take the next step, mainly booking an appointment with a human sales representative.

This structured approach ensures that the AI agent effectively engages potential customers, addresses their needs, and persuades them to take action, thereby enhancing the overall sales process.
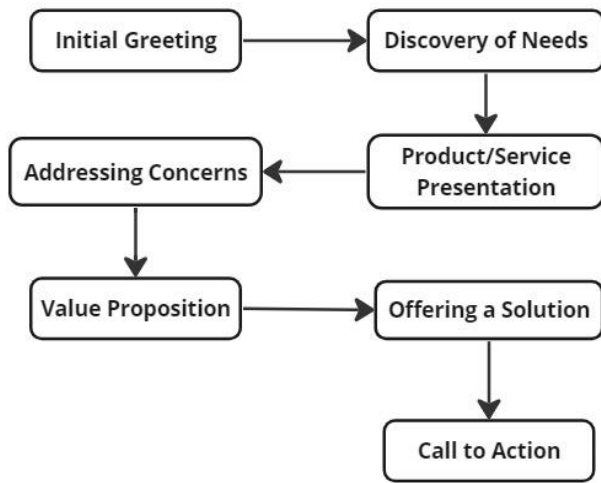

Fig. 2.   AI Agent Call Structure Diagram

IV.   IMPLEMENTATION AND RESULTS

The implementation of the AI-based call agent was achieved using various advanced technologies. The system components and their respective alternatives are listed in Table 1, with the chosen technologies highlighted for their performance and seamless integration.

We selected Twilio for the telephony API due to its reliable performance and ease of integration. Deepgram was chosen for the speech-to-text model because of its accuracy and robustness in handling various accents and speech patterns. ElevenLabs was preferred for the text-to-speech

model due to its natural-sounding voice synthesis. Finally, GPT-3.5 was selected as the LLM for its advanced conversational capabilities and ability to generate contextually relevant responses.

TABLE I.   SYSTEM COMPONENTS ALTERNATIVES

System Component	Alternatives
Telephony API	**Twilio**, Vonage, Plivo
Speech-To-Text Model	**Deepgram**, AssemblyAI, Whisper by OpenAI
Text-To-Speech Model	**ElevenLabs**, PlayHT, Deepgram
LLM	**GPT-3.5**, mixtral-8x7b, Llama3-8b-8192

A.   Latency Evaluation

Our latency tests showed that the AI-based call agent had an average response time of 1.7 seconds across various query complexities. For example, when initiating a conversation with a simple greeting, the response time was consistently under 1 second. When answering straightforward questions like "How much does it cost?", the average response time was around 1.4 seconds. For more complex inquiries, such as "Can you explain how this product integrates with our existing systems?", the response time averaged 2.5 seconds. The maximum response time recorded during peak testing periods was 2.8 seconds, ensuring a natural and engaging interaction flow. The system also handled concurrent calls smoothly, with no significant delays even under increased call volumes.

Figure 3 illustrates the average latency of each component in the AI-based call agent system, demonstrating the efficient performance of the chosen technologies. The latency for all system components contribute to the overall response time, which adds up to the average of 1.7 seconds.
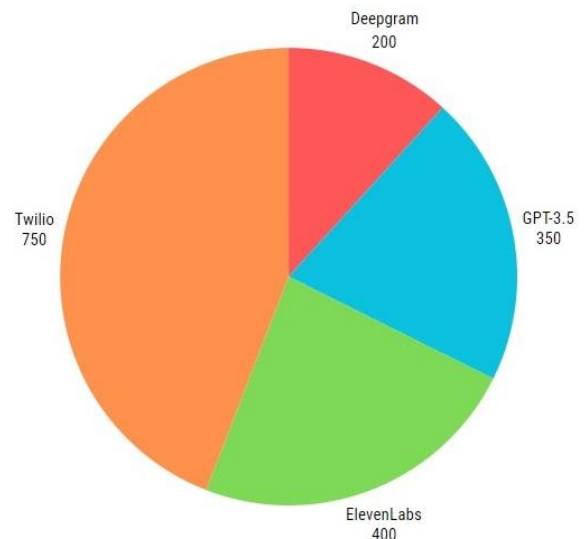

Fig. 3.   Average Latency of System Components in Milliseconds

B.   Conversion Rate

To evaluate the conversion rate, we conducted 25 test cold calls with the objective of introducing a product or

service and convincing the recipient to book an appointment with a human sales representative. The AI agent successfully converted 2 out of the 25 calls, resulting in a conversion rate of 8%. This performance is notable given that typical industry cold call conversion rates range from 2-5%. The AI agent's ability to sound realistic, incorporating casual language and humor, contributed significantly to engaging customers effectively.

C. Customer Satisfaction

Post-call surveys revealed positive feedback regarding the AI agent's performance. Customers rated the AI's realism highly, with 70% agreeing that it sounded almost human. The agent's friendliness received an average score of 4.8 out of 5, while clarity of information scored a perfect 5.0. The response time was praised, averaging 4.2 out of 5. When compared to human interactions, 70% of respondents felt that the AI agent provided a comparable or superior experience, highlighting its consistency and efficiency.

V. FUTURE WORK

Our work offers a strong foundation for revolutionizing customer interactions, and there are numerous opportunities to enhance and expand its capabilities. One promising direction is tailoring the system for industry-specific applications. By fine-tuning the underlying LLM using domain-specific data, the agent can become adept at handling specialized terminology and meeting regulatory requirements in fields such as healthcare, finance, and real estate. These enhancements will enable the agent to deliver precise and contextually relevant interactions, making it a valuable tool across various industries.

Advancements in the agent's speech-to-text and text-to-speech components present another exciting avenue for development. Improvements in handling diverse accents, speech patterns, and challenging environmental conditions will ensure the system performs consistently in real-world scenarios. Such refinements will enhance the robustness and reliability of the agent, providing seamless interactions regardless of external factors.

Building on these advancements, scalability and enriched functionality will further expand the agent's potential. Efforts to optimize performance will enable it to handle high call volumes efficiently, ensuring quick response times and maintaining interaction quality. This scalability will support deployment in large-scale environments like call centers, allowing the agent to serve a wider audience without compromising reliability. Additionally, new features such as real-time sentiment analysis, adaptive dialogue flows, and seamless CRM integrations will enhance its capabilities. These improvements will enable the agent to deliver personalized and impactful customer experiences, making it an even more versatile and valuable tool across industries.

VI. CONCLUSION

The AI-based call agent has demonstrated great potential in revolutionizing sales processes by delivering consistent, efficient, and scalable customer interactions. With an average response time of 1.7 seconds, an 8% conversion rate, and strong customer satisfaction ratings, the system has proven to exceed industry averages, showcasing its ability to enhance engagement and drive sales effectively. These results affirm the potential of AI-driven solutions to address key challenges faced by traditional sales methods, such as operational costs and variability in performance.

The findings underscore the agent's capacity to operate as a transformative tool for businesses, leveraging advancements in NLP and voice technologies to deliver personalized and impactful customer experiences. By refining its capabilities and expanding its applications, this system has the potential to further modernize sales strategies and redefine industry standards.

REFERENCES

[1] C. Roberts and F. Alpert, "Total customer engagement: designing and aligning key strategic elements to achieve growth," *Journal of Product & Brand Management*, vol. 19, no. 3, pp. 198–209, Jun. 2010, doi: https://doi.org/10.1108/10610421011046175.

[2] Z. Zhang, B. Li, and L. Liu, "The impact of AI-based conversational agent on the firms' operational performance: Empirical evidence from a call center," Applied Artificial Intelligence, vol. 37, no. 1, Dec. 2022, doi: 10.1080/08839514.2022.2157592.

[3] L. Wang, N. Huang, Y. Hong, L. Liu, X. Guo, and G. Chen, "Voice-based AI in call center customer service: A natural field experiment," Production and Operations Management, vol. 32, no. 4, pp. 1002–1018, Apr. 2023, doi: 10.1111/poms.13953.

[4] S. Choudhary, N. Kaushik, B. Sivathanu, and N. P. Rana, "Assessing factors influencing customers' adoption of AI-Based voice assistants," Journal of Computer Information Systems, pp. 1–18, Feb. 2024, doi: 10.1080/08874417.2024.2312858.

[5] K. Akdim and L. V. Casaló, "Perceived value of AI-based recommendations service: the case of voice assistants," Service Business, vol. 17, no. 1, pp. 81–112, Feb. 2023, doi: 10.1007/s11628-023-00527-x.

[6] C. Sun, Z. Shi, X. Liu, A. Ghose, X. Li, and F. Xiong, "The effect of voice AI on consumer purchase and search behavior," Social Science Research Network, Jan. 2019, doi: 10.2139/ssrn.3480877.

[7] A. Mari, R. Algesheimer, and N. Outi, "AI-based voice assistants for digital marketing: preparing for voice marketing and commerce," in Contemporary Issues in Digital Marketing, 2021. [Online]. Available: https://www.zora.uzh.ch/id/eprint/208002/

[8] S. Jusoh, "Intelligent conversational agent for online sales," 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Jun. 2018, doi: 10.1109/ecai.2018.8679045.

[9] A. Mari, A. Mandelli, and R. Algesheimer, "The Evolution of Marketing in the context of Voice Commerce: A Managerial perspective," in Lecture notes in computer science, 2020, pp. 405–425. doi: 10.1007/978-3-030-50341-3_32.

[10] S. H. Aldulaimi, M. M. Abdeldayem, B. M. Mowafak, and M. M. Abdulaziz, "Experimental Perspective of Artificial Intelligence Technology in Human Resources Management," Applications of Artificial Intelligence in Business, Education and Healthcare, pp. 487–511, 2021, doi: https://doi.org/10.1007/978-3-030-72080-3_26.