ISSN: 2508-7894
KJAI website: http://www.kjai.or.kr
doi: http://dx.doi.org/10.24225/kjai.2023.11.4.9

Research on Developing a Conversational AI Callbot Solution for Medical Counselling

Won Ro LEE[1], Jeong Hyon CHOI[2], Min Soo KANG[3]

Received: November 09, 2023. Revised: December 04, 2023. Accepted: December 08, 2023.

Abstract

In this study, we explored the potential of integrating interactive AI callbot technology into the medical consultation domain as part of a broader service development initiative. Aimed at enhancing patient satisfaction, the AI callbot was designed to efficiently address queries from hospitals' primary users, especially the elderly and those using phone services. By incorporating an AI-driven callbot into the hospital's customer service center, routine tasks such as appointment modifications and cancellations were efficiently managed by the AI Callbot Agent. On the other hand, tasks requiring more detailed attention or specialization were addressed by Human Agents, ensuring a balanced and collaborative approach. The deep learning model for voice recognition for this study was based on the Transformer model and fine-tuned to fit the medical field using a pre-trained model. Existing recording files were converted into learning data to perform SSL(self-supervised learning) Model was implemented. The ANN (Artificial neural network) neural network model was used to analyze voice signals and interpret them as text, and after actual application, the intent was enriched through reinforcement learning to continuously improve accuracy. In the case of TTS(Text To Speech), the Transformer model was applied to Text Analysis, Acoustic model, and Vocoder, and Google's Natural Language API was applied to recognize intent. As the research progresses, there are challenges to solve, such as interconnection issues between various EMR providers, problems with doctor's time slots, problems with two or more hospital appointments, and problems with patient use. However, there are specialized problems that are easy to make reservations. Implementation of the callbot service in hospitals appears to be applicable immediately.

Keywords : Shared bikes, Demand forecasts, Linear regression, Machine learning, AI

JEL Classification Code: E44, F31, F37, G15

1. Introduction

Speech recognition, the core of conversational AI systems, is making leaps and bounds thanks to deep learning, one of the subfields of artificial intelligence. Deep learning generally refers to the training of "deep" artificial neural networks (ANNs). Advances in deep learning techniques have made it possible to address the limitations of traditional statistical and rule-based natural language understanding

* This work was supported by the research grant of the KODISA Scholarship Foundation in 2023
1 First Author. NascaLab. Korea, Email: nascalab@gmail.com
2 Second Author. Director, LG Uplus corporation, Korea, Email: nevergiveup@lguplus.co.kr
3Corresponding Author. Professor, Dept. of Medical Bigdata, Eulji Univ. Korea, Email: mskang@eulji.ac.kr

(NLU) methods and machine learning approaches to NLU. Traditional NLU methods of the past few decades have relied heavily on manual labor. For this reason, they were time-consuming to extract and contained various imperfections.

However, in recent years, artificial neural networks (ANNs) based on dense vector representations have demonstrated superior performance on a variety of natural language tasks, largely due to the success of word embeddings and deep learning techniques. For example, Google's Assistant, which is ranked #1 in the 2022 Gartner® Magic Quadrant for Cloud AI Developer Services, uses deep learning techniques to learn from speech data to reduce errors and improve accuracy in speech recognition, bringing its speech recognition accuracy from less than 80% in 2013 to more than 95%, a level of accuracy that is comparable to that of a human (Gartner, 2022).

Figure 1: 2022 Gartner Report Magic Quadrant for Cloud A I Developer Services

Improvements in speech recognition accuracy are leading to advances in conversational AI-enabled services and rapidly defining the overall conversational AI market. Conversational AI systems are becoming increasingly sophisticated and are being applied across a wide range of industries for a variety of benefits, including 24/7 customer support, labor cost savings due to reduced human dependency, and automation of repetitive processes.

There are two main types of interactive AI counseling systems: messenger-based chatbot and voice-based callbot (voicebots). A chatbot is a combination of chat and robot, and it refers to an interactive system that interacts with people based on artificial intelligence (Park, 2020). Like a chatbot, a callbot is a service that combines speech recognition technology and speech synthesis technology, and can be summarized as a "talking chatbot". In other words, if a chatbot is a test-based communication service, a callbot is a service that communicates by voice as if you were talking to an agent.

Chatbot have the advantages of being instantaneous, allowing users to send questions and check answers at their own time, personalized communication, and quick loading on messenger platforms such as Facebook and KakaoTalk. However, the limitations of chatbot stem from the limitations of the creation model and search model for creating chatbot. Since they are implemented based on scenarios, they often rely on a simple search model, and since humans have to design and intervene, they often fail to respond appropriately. Callbot, on the other hand, are voice-based and can provide the fastest communication, instantly, without running any software, and can be guided to understand what the customer wants and select an efficient response, making them accessible and convenient. They can also successfully perform multiple tasks simultaneously by continuously improving the customer experience. On the other hand, the limitations of callbot include problems with speech recognition due to different dialects and pronunciations, problems with recognizing domain-specific speech errors, and problems with distinguishig between noise and user voice. Nevertheless, voice is the fastest form of communication in human history, and the fact that it is available over the phone and can be improved through continuous learning makes it an attractive tool for companies operating phone-based contact centers. Currently, the global conversational AI market is dominated by chatbot. This is because the limitations of callbot have not been able to overcome the advantages of chatbot.

However, if we look at the medical consultation field, which is mainly used by the elderly, despite the introduction of chatbot services, consultations are mainly conducted via telephone. In the current situation, when applying chatbot to hospitals, dermatology, plastic surgery, and pediatrics, which are used by relatively young people, the consultation system through chatbot is effective, but in the case of most hospitals where the main users are elderly people, there are limitations in the actual use of chatbot due to the phenomenon that users are vulnerable to digital devices and prefer to communicate through phone rather than messenger.

Building upon the observation that a significant portion of the elderly population frequents medical institutions, our research aimed to introduce a conversational AI call bot service for these healthcare settings. According to the '2022 Health Insurance Key Statistics' by the National Health Insurance Service, the combined medical expenses, encompassing both the state's and individuals' shares, totaled KRW 102.4277 trillion. Out of this sum, the medical expenditure for those aged 65 and over amounted to KRW 44.1187 trillion, marking an 8.6% increase from the previous year and representing 43.1% of the entire medical cost (National Health Insurance Service, 2022). With South Korea's rapid aging trajectory, it's projected that the elderly's medical expenses will soon surpass 50% of the total in the coming years. This implies that a majority of hospital attendees will be from the elderly demographic. In such a scenario, we recognized the pressing need for an interactive

AI call bot to facilitate hospital consultations, especially considering the elderly's inclination towards phone-based interactions.

2. Related Research

While providing IPCC (Internet Protocol Contact Center) solutions and CRM software to about 120 hospitals, the average response rate of hospitals is around 80-85%, and on average, more than 15% of calls to hospitals are not connected to the customer service center. It was found that It was found that even in the case of connected calls, about 30% are related to simple consultations, such as location information, treatment time information, inquiries about medical certificates and confirmations, and confirmation, change, or cancellation of treatment reservations for returning patients.

As a way to reduce simple consultation work, based on linking with EMR (Electronic Medical Record), treatment time and location information is automatically provided through BMS (Back Massage Service) to patients who call the hospital for the first time, or location information is provided on ARS. If you request information, treatment time information, proof of registration, etc., a text message is automatically sent.

Going further, we use the IPCC (Internet Protocol Contact Center) solution and TTS (Text To Speech) engine to link reservation information on the hospital's EMR so that patients can confirm, change, or cancel their medical reservations through IPCC's ARS using DTMF on their phones. It was implemented to process it as a signal value.

Furthermore, to prevent NOSHOW, IPCC's ACS (Auto Calling Service) function was used to automatically call patients to confirm, change, or cancel medical appointments 3 days before the appointment date instead of sending a text message, thereby reducing NOSHOW by about 3%. .

Through related research, it was possible to resolve about 10% of unconnected calls and simple inquiry calls through an intelligent ARS system using TTS. For this reason, it is currently being used in many hospitals.

In this study, we went further on existing related research and advanced it through conversational AI callbot to improve the hospital's response rate and to handle more simple consultation cases, resulting in 30% of simple inquiries rather than medical consultations and 30% of unconnected calls. By processing 15%, we started with a goal of increasing the work efficiency of customer center agents and the work environment.

In other words, simple inquiries and confirmation services are handled by AI Agents, and parts requiring consultation are handled by Human Agents, with the goal of reducing costs through efficient resource management and human resources management at hospitals and improving productivity through optimization of resource operation. It went ahead.

3. Experiment

The final goal of this study was to improve the hospital customer service response rate and improve the hospital NOSHOW as shown in Table 1

Table 1: The end goal of a medical advice callbot's service

Goals	Achieve	Detailed goals
Improve consultation quality	Improve hospital customer service response rates	Hospital customer center responds to 30% of simple inquiries and simple reservations with AI voice consultation system
		Hospital agents available for off-hours, holidays, and simple consultations
		90%+ hospital customer response rate
Improving appointment management	Improve hospital NOSHOW	Hospital NO SHOW Improvement of 3 Hospital Average NO SHOW 15%_Fair Value Data

The speech recognition (TTS), speech synthesis (STT), and dialogue engines were developed using engines secured from NASCAR Labs as part of the "2020 AI Voucher Support Project" of the Korea Information and Communication Industry Promotion Agency, and hospitals interested in applying the service were selected and conducted as a pilot service. As shown in Figure 5 below, we first collected customer requirements and created FAQs, separated the recorded data in the customer's system into listeners and speakers, converted it into text, and de-identified and labeled personal information to obtain about 600 minutes of learning data. This data was used for the first round of AI training. The conversation engine was used to recognize conversation intent, manage conversation flow, and manage conversation context, and dialect, honorifics, and abbreviations were learned by organizing the corresponding intent. The IVR scenario of the IPCC solution was configured by separating the connection between the callbot agent and the human agent, and if the callbot agent does not recognize more than twice, it is connected to the human agent during business hours and handled as a callback at the end of the business day. The call bot reads patient information and reservation information linked to the hospital's EMR, and the reservation function is implemented by establishing a development process so that the call bot

speaks through TTS, and the matters corresponding to general guidance are learned by enriching the intent.

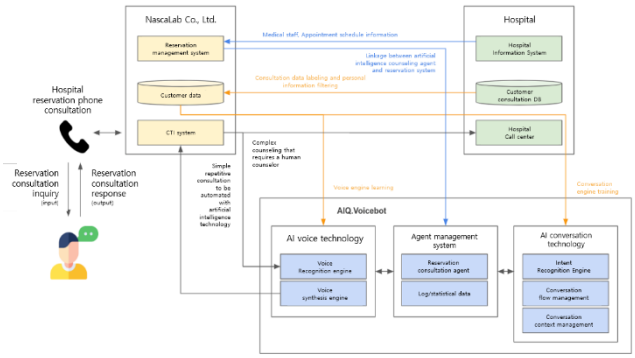

Figure 2: Process to development of medical consultation Callbot

3.1. STT Training

We de-identified and labeled 600 hours of personal information for the medical domain by separating listeners and speakers from the recorded data held by the customer, and applied it to the hospital STT project and conducted STT learning in parallel. Through this, we were able to increase the STT recognition rate for the medical domain from 75% to 93%.

Due to problems such as slow response speed and inaccurate pronunciation of the existing TTS engine, the TTS engine was purchased and applied from a well-known domestic company, and a project was created and linked between STT and TTS.

Based on the FAQ information organized by the hospital , information such as parking information, COVID-19 information, clinic time and location information, and su rrogate care was provided to the TTS through the trigger int ent, and it was implemented as shown in Figure 3 to go dire ctly to the corresponding dialog when the intent is expressed.

Figure 3: Chat-Bot Scenario

If there is a reservation-related item in the intent, the reservation information of the patient is checked in EMR through the patient's incoming phone number and date of birth, and the reservation information is informed. Since EMR is personal sensitive medical information and the hospital's timekeeping soft, only the reservation date and doctor information are linked, not the entire information. To connect with EMR, a mediation server was installed inside the hospital and set as a proxy bridge between EMR DB and CTI DB. In addition, due to the situation that the time slot management for each medical department and medical staff is managed separately in units of 5 minutes, 10 minutes, 15 minutes, and 30 minutes, the EMR reservation system required the cooperation and consent of each outpatient to allow the callbot to make and change appointments only in 30-minute increments.

The intention of the first hospital was to reduce the recognition rate due to the difference in expression methods between written expressions and actual customers using various colloquial expressions and dialects, so to compensate for this, we linked the Call-Bot directly to the IPCC and called the hospital staff before applying the service, so that even if the AI does not respond, the general public can ask and talk about what they want to know about the hospital or inquire about during hospital business without bias. Through the above pre-open, we were able to obtain interactive intent in the form of questions asked by the public, which was immediately linked to the corresponding guidance trigger. After that, we proceeded with the training as shown in Figure 4 to improve the accuracy.

Figure 4: Collect Intent and expend dataset

This allowed us to add new scenarios that differed from the FAQs initially provided by the hospital, such as health card inquiries, pharmacy guidance, proxy prescriptions, proxy verification procedures, infant and child check-ups, out-of-school check-ups for teens, and responses to questions about room costs and non-paid medical expenses that are not often asked when talking to people.

4. Result

After more than a year of information research and development, Ver1 is now ready for piloting and is still being tested. There are many shortcomings and a long way to go for the service we are aiming for. First, the success rate of voice recognition is currently close to 93%, but the recognition rate for background noise, inaccurate pronunciation, and dialect is still low. To overcome this, we set the system to connect to a human agent or callback if the callbot agent does not recognize the caller more than twice. It is necessary to find a way to remove noise during voice input. Second, there is the problem of applying medical staff time slots. The time slot of medical staff in hospitals varies from 5 minutes, 10 minutes, and 30 minutes depending on the department and doctor, and there are cases where there are consultations with other departments, and there are cases where there are pre-treatment examinations or post-treatment examinations. Communicating with callbot and patients by linking this part with EMR is different from simply making a reservation for a restaurant or ordering a certain food according to a set time slot. We need to think more about this and find a solution, and as a way to solve this, we are currently applying a method that connects to a human agent if there are more than two departments or tests in the future medical treatment at that time. Third, continuous reinforcement learning is necessary for the sophistication of the current system. Until a certain level of service level is achieved, continuous investment of time and money is required.

5. Conclusions

The market for conversational AI using artificial intelligence, one of the important technologies of the Fourth Industrial Revolution, continues to grow and is expected to develop further with the development of speech recognition technology and the enhancement of the performance of CHAT GPT. However, as we have seen in this study, there are still problems that need to be solved before callbot can be applied to the field of hospital consultation to provide a smooth service. In addition, in order to provide services to the elderly, it is necessary to fight against time and cost to solve various worries and problems. However, we saw the possibility of service in the process, and we also saw the possibility of service in specialized hospitals where medical appointments are not complicated. As part of the medical consultation customer center solution, we look forward to the era of a customer center where callbot create a callbot agent avatar that fits the image of the hospital, consult based on artificial intelligence, and provide cutting-edge customer service through digital assistants. We hope that the quality of medical consultation will be enhanced through mutual cooperation between callbot agents and human agents.

References

Cha, K., Kim, D., & Lee, Y. (2018). A Study on Voice-based Conversational User Interface Design for Maximizing User Experience. *KITS*, *Integrated Academic Conference on Management-Related Studies,* 2018.08. http://www.itservice.or.kr/data02.html?bmode=read&bid=data1&id_no=364&l=1

Choi, S. (2015). Health Insurance Medical Expenses Expenditure Status and Implications for Health Insurance Financial Management*. Korea Institute of Public Finance.* https://www.kipf.re.kr/cmm/fms/FileDown.do;jsessionid=6CD5010D3B1F4203F49F4095B60CC9EE?atchFileId=FILE_000000000007255&fileSn=0

Gartner. (2022). Magic Quadrant for Cloud AI Developer Services. Gartner Report.

Hwang, M. (2020). Effects of the Characteristics of Experience of the Chatbot on the Hospital Image. *Gachon University.*

INNOPOLIS FOUNDATION. (2021). Conversational Artificial Intelligence (AI) Market. *Global Market Trend Report.* https://www.innopolis.or.kr

LIVY Co., Ltd. (2018). Development of Conversational Solution for Intelligent Chat Services Based on Pragmatic and Context Analysis of Dialogues. *Ministry of Science and ICT*.

LGCNS blog.(2018). NLU, a machine that understands human language. https://www.lgcns.com/blog/cns-tech/ai-data/2343/

National Health Insurance Service. (2022). 2022 Major Health Insurance Statistics.

NascaLab Co., Ltd. Additional services: BMS, ACS, medical appointment confirmation, change, cancellation, NOSHOW management. http://www.nascalab.com/

Park,Arum.(2020) Application of AI based Chatbot Technology in the Industry. *KSCI. 25*(7), 17-25, July 2020 https://doi.org/10.9708/jksci.2020.25.07.017

Park, P. W. (2020). Text-CNN Based Intent Classification Method for Automatic Input of Intent Sentences in Chatbot. *Journal of KIIT*, *18*(1), 19-25. https://doi.org/10.14801/jkiit.2020.18.1.19

Ryu, Ki-Dong (2019). “Research on artificial intelligence-based contact center system”, *Korean doctoral thesis*, Seoul University of Science and Technology, Seoul

Oh, Y. R., Park, K., & Park, J. G. (2022). Fast offline transformer-based end-to-end automatic speech recognition for real-world applications. *ETRI Journal*, *44*(3), 476-490.

Nam, K. (2019). A study on processing of speech recognition Korean words. The Journal of the Convergence on Culture Technology, Changhan Oh et al. / *Phonetics and Speech Sciences* Vol.15 No.3 (2023) 75-82 81 *5*(4), 407-412