

Knowledge and Language-processing Technology that Supports and Substitutes Customer-contact Work

Kyosuke Nishida, Kuniko Saito, Tetsuo Amakasu, Kazuyuki Iso, and Shuichi Nishioka

Abstract

At NTT Media Intelligence Laboratories, using natural-language-processing technology cultivated over many years as one of our core competences, we are researching and developing knowledge and language-processing technology that contributes to improving the productivity of contact centers and offices. In this article, a language model, document-summarization technology, and interview-support technology are introduced.

Keywords: AI, document summarization, response analysis

1. Introduction

At NTT Media Intelligence Laboratories, we have researched and developed knowledge and language-processing technology for contact centers for analyzing documents (such as work manuals and frequently asked questions) and presenting appropriate documents to agents handling customers. Not only in contact centers but also in offices, it is necessary to further improve productivity of agents and employees. In response to this necessity, we are developing technology for understanding and generating large-scale documents and various responses. We first explain a language model for handling documents and document-summarization technology using this language model then describe interview-support technology for analyzing responses between a customer and agent.

2. Development of natural-language understanding by using a language model (BERT)

It has been considered difficult for artificial intelligence (AI) to understand human language. However,

with the advent of Bidirectional Encoder Representations from Transformers (BERT) [1], announced by Google in October 2018, the research and development (R&D) of natural-language understanding has undergone a large paradigm shift. For example, regarding the task called *machine reading* [2], which requires reading comprehension to understand the content of text and answer questions, it has been reported that AI using BERT greatly exceeded the response score of a human. Performance of natural-language-processing tasks other than machine reading has also improved significantly, and language models are attracting attention as a basic technology for giving AI the ability of language comprehension.

A language model estimates the plausibility of sentences (**Fig. 1**). For example, many people may feel that “cake” is more natural than “egg” as the word “X” in the sentence “Today is my birthday, so I ate [X].” Also, it seems natural that the two sentences “It’s a fine day today.” and “It’s a good day to do the laundry.” appear consecutively. BERT learns in advance such fill-in-the-blank problems (word prediction) and relationships between two consecutive sentences (prediction of the next sentence) on the basis of a large set of texts such as all Wikipedia

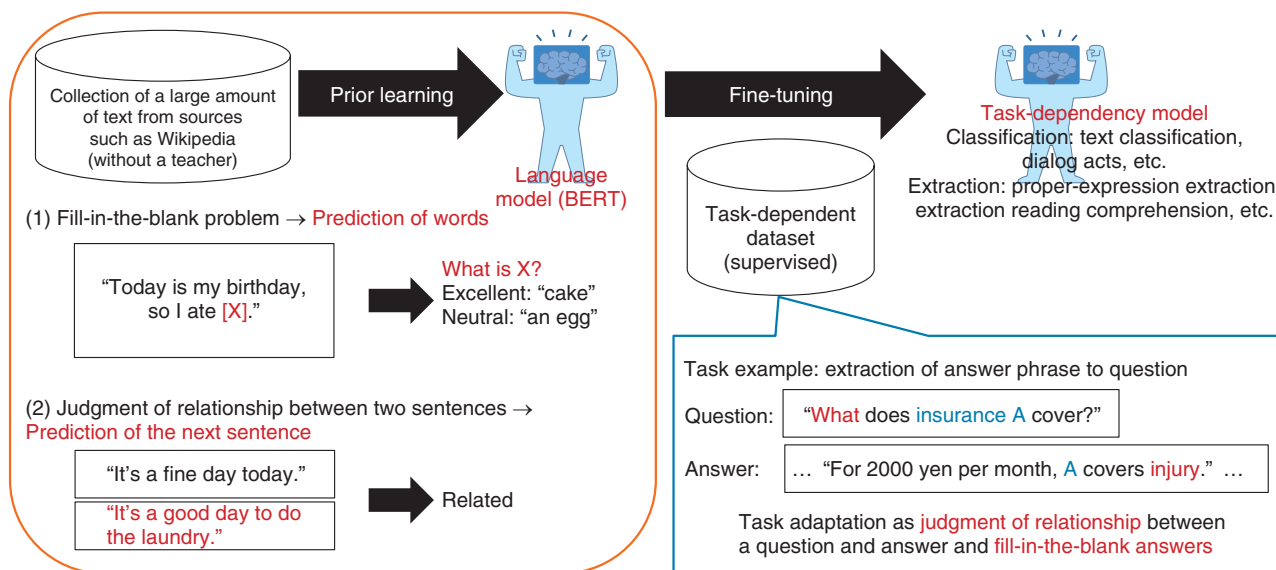


Fig. 1. Language model BERT.

sentences. By learning (fine-tuning) with various task-dependent datasets, the language model (BERT) obtained in this manner can be applied to various tasks (such as classifying texts by genre and extracting phrases that will answer questions) and achieve high performance even when a large amount of learning data is not available for the applied task.

BERT has had a great impact on natural-language processing, and language models are still being researched and constructed all over the world. At NTT Media Intelligence Laboratories, while collecting a large amount of Japanese text data and creating a Japanese-language BERT, we are researching technologies that use language models for tasks such as summarizing documents [3, 4], retrieving documents [5], and answering questions [6, 7]. In each case, high performance is achieved by not only simply applying BERT but also using the knowledge we have gained through our previous research on natural-language processing and deep learning. Moreover, by investigating the characteristics and internal operation of BERT [8], we are researching with the aim of constructing our own language model that remedies the shortcomings of BERT.

3. Document-summarization technology for summarizing documents by specifying length

Document-summarization technology is introduced as a representative example of a technology

that uses the language model described above. Document summarization has been grappled with for many years. For customer-contact work such as at contact centers, if AI returns a long sentence as a result of searching previous calls and answering a customer's question, it will be difficult for the customer to read such a long sentence; accordingly, it is desirable to appropriately adjust the length of the sentence.

Given the above-mentioned issue, NTT Media Intelligence Laboratories developed a document-summarization technology that can control sentence length by using a neural network [3]. This technology consists of a combination of an extractive model (which identifies important points in a sentence) and abstractive model (which generates a summary sentence from the original sentence). The extractive model learns on the basis of the language model. By controlling the number of important words output by the extractive model according to the specified sentence length and by generating a summary sentence based on both the important words and the original sentence, it was possible to establish a technology that can control sentence length and summarize with high accuracy.

This document-summarization technology is used as the core engine of the COTOHA™ API summarization function developed by NTT Communications [9]. COTOHA Summarize offers a service for outputting a summary text when a document is input and provides our customers a free tool for generating

summary texts of sites browsed with a web browser. We plan to further develop this technology for the NTT Group.

We will continue to advance R&D to improve this summarization technology, which includes not only document summarization but also summarization that targets dialogue and summarization that allows the viewpoint and keywords of the summarization to be specified externally. To further strengthen the competitiveness of our language model, we plan to continue technological development while incorporating the results of our latest research on language processing. Such work will include scaling up the model, constructing a generative language model for generating more-natural sentences, and establishing summarization technology based on that model [4].

4. Technology for using knowledge obtained from contact-center calls

4.1 Challenges in supporting customer contacts

NTT Media Intelligence Laboratories developed an automatic knowledge assistance system [10]. This is a technology that automatically retrieves and presents documents according to the content of a conversation with a customer. By supporting the agent with knowledge and responding promptly to customers with appropriate information, it became possible to improve relationships with customers.

However, contact centers are required to play a new role as the business style changes due to the digital transformation (DX) of offices and the spread of the novel coronavirus. One of these roles is as a sales method called *inside sales*. Inside sales is explained as follows. In contrast to the conventional sales method, namely, a full-time salesperson conducts face-to-face sales to understand customer needs and conclude business negotiations, inside sales is based on (i) maintaining continuous communication with the other party while gathering information (such as understanding needs) that triggers business negotiations via telephone or web conference and (ii) dispatching a salesperson when the possibility of receiving an order or contract increases. The number of contact centers that handle customers by using inside sales is increasing. The challenges facing contact centers playing this role are as follows.

- Further improvement of agent or productivity: When creating a report after finishing a call, it is necessary to support the extraction of important information in the response from business negotiations in which the flow of conversation is

complicated and the topics are diverse.

- Improvement of productivity of sales-information analysis: From the standpoint of the agents' supervisor, it is necessary to understand and analyze the likelihood of making contracts, the tendency of customer needs, and other conversational tendencies included in business negotiations between each agent and customers. Planning and agent operations are then improved on the basis of the analyzed information. It is thus necessary to support such analysis and improvement work.

4.2 Interview-support technology

To address the above-mentioned challenges, we developed interview-support technology. This technology consists of two elements (Fig. 2). The first specifies a series of responses to the customer for each section of the topic. The second extracts important utterances such as questions, answers, and explanations from the sections specified by the former.

A characteristic of a business conversation that draws out issues and requests concerning a customer is that the topic continuously changes according to the answers given by the customer. Since conventional technology cannot cope with such a dynamic situation, we developed a means of firmly determining the point at which the topic changes. This is the first feature of the interview-support technology. Sections of the same topic are specified by machine learning using characteristic expressions at the point of change of the topic appearing in a conversation, and the type of topic is acquired by machine learning using characteristic expressions in the topic. The second feature of this technology is that the utterances of questions and answers of agents and customers for each topic are extracted by machine learning using characteristic expressions included in important utterances such as questions and answers.

By using the text of a call that is divided into sections for each topic by using the interview-support technology, after the call, the agent can quickly identify the points at which he/she heard about the customer's needs and budget and create a report based on that information. Moreover, by collecting such information from many calls, it becomes possible to support the analysis of sales information. We aim to apply this technology to responses via voice but also via online chat (text messages), which has been increasing in popularity.

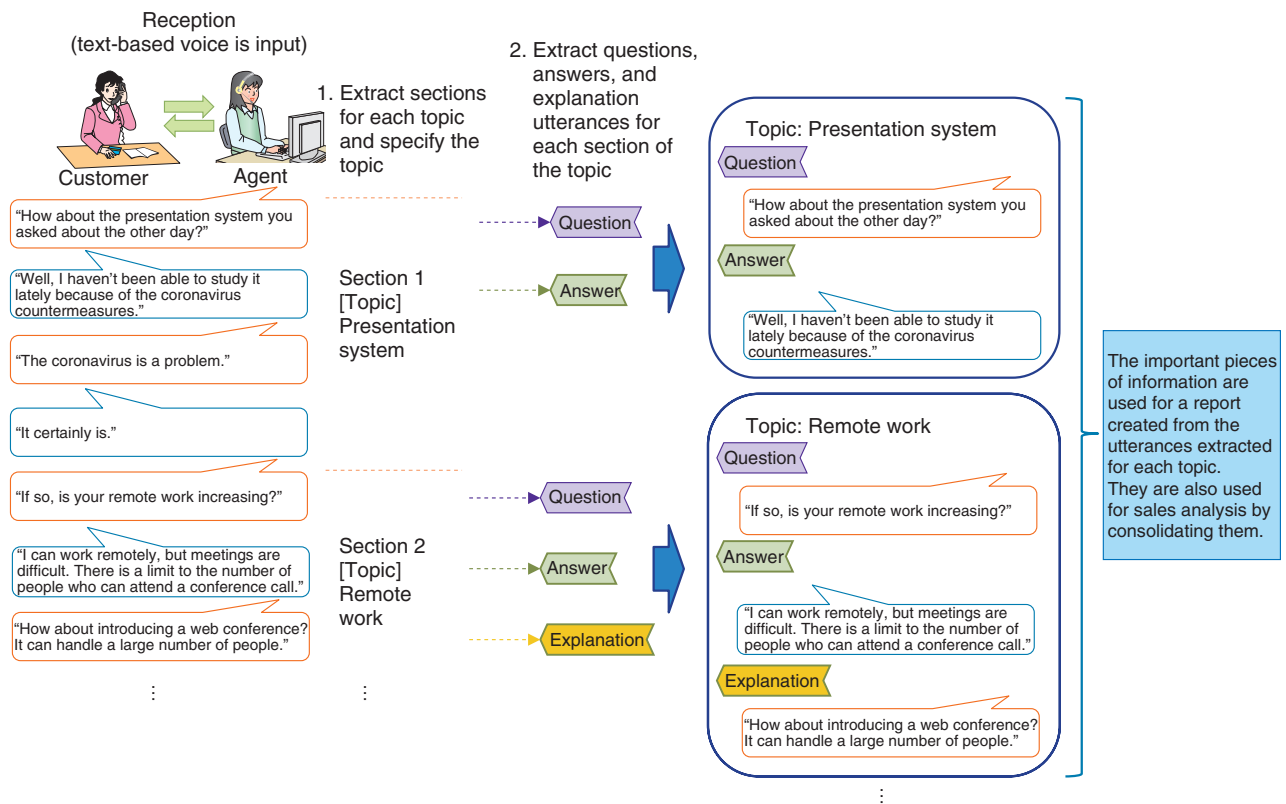


Fig. 2. Interview-support technology.

5. Concluding remarks

We introduced our language model, document-summarization technology, and interview-support technology. Going forward, we plan to work on two technologies for understanding and generating large-scale documents and various responses: (i) a technology that can read various document layouts and search for necessary information at high speed and high accuracy and (ii) a technology that supports strategic conversations by grasping the details of a conversation and extracting needs that customers may have difficulty noticing themselves.

References

[1] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," Proc. of NAACL-HLT 2019, Vol. 1, pp. 4171–4186, Minneapolis, Minnesota, USA, June 2019.

[2] K. Nishida, I. Saito, A. Otsuka, K. Nishida, N. Nomoto, and H. Asano, "Toward Natural Language Understanding by Machine Reading Comprehension," NTT Technical Review, Vol. 17, No. 9, pp. 9–14, Sept. 2019. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201909fa3.html>

[3] I. Saito, K. Nishida, K. Nishida, A. Otsuka, H. Asano, J. Tomita, H. Shindo, and Y. Matsumoto, "Length-controllable Abstractive Summarization by Guiding with Summary Prototype," arXiv, 2001.07331, Jan. 2020.

[4] I. Saito, K. Nishida, K. Nishida, and J. Tomita, "Abstractive Summarization with Combination of Pre-trained Sequence-to-Sequence and Saliency Models," arXiv, 2003.13028, Mar. 2020.

[5] T. Hasegawa, K. Nishida, S. Kaku, and J. Tomita, "Sparse Contextual Sentence Representation for Fast Information Retrieval," Proc. of the 34th Japanese Society of Artificial Intelligence (JSAI 2020), Mar. 2020 (in Japanese).

[6] K. Nishida, K. Nishida, I. Saito, H. Asano, and J. Tomita, "Unsupervised Domain Adaptation of Language Models for Reading Comprehension," Proc. of the 12th Conference on Language Resources and Evaluation (LREC 2020), pp. 5392–5399, May 2020.

[7] K. Nishida, K. Nishida, I. Saito, H. Asano, and J. Tomita, "Explainable Machine Reading Comprehension," Proc. of the 26th Annual Meeting of the Association for Natural Language Processing (NLP 2020), pp. 1–19, Mar. 2020 (in Japanese).

[8] Y. Ohsugi, I. Saito, K. Nishida, H. Asano, and J. Tomita, "How Do Masked Language Models Perform When the Input Sequence Length Changes?," Proc. of JSAI 2020, Mar. 2020.

[9] COTOHA™ API, <https://api.ce-cotoha.com/contents/about.html>

[10] T. Hasegawa, Y. Sekiguchi, S. Yamada, and M. Tamoto, "Automatic Knowledge Assistance System Supporting Operator Responses," NTT Technical Review, Vol. 17, No. 9, pp. 15–18, 2019. <https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201909fa4.html>



Kyosuke Nishida

Distinguished Researcher, Knowledge Media Project, NTT Media Intelligence Laboratories.

He received a B.E., MIS, and Ph.D. in information science and technology from Hokkaido University in 2004, 2006, and 2008 and joined NTT in 2009. He was a chief engineer at NTT Resonant from 2013 to 2015. He received the Yamashita SIG Research Award from the Information Processing Society of Japan (IPSJ) in 2015 and the Kambayashi Young Researcher Award from the Database Society of Japan (DBSJ) in 2017. His current interests include natural-language processing and AI. He is a member of the Association for Computing Machinery (ACM), IPSJ, the Association for Natural Language Processing (NLP), the Japanese Society for Artificial Intelligence (JSAI), the Institute of Electronics, Information and Communication Engineers (IEICE), and DBSJ.



Kuniko Saito

Senior Research Engineer, Supervisor, NTT Media Intelligence Laboratories.

She received a B.E. and M.E. in chemistry from the University of Tokyo in 1996 and 1998. She joined NTT Information and Communication Systems Laboratories in 1998. Her research focuses on natural-language processing and AI. She is a member of NLP and IPSJ.



Tetsuo Amakasu

Senior Research Engineer, Supervisor, Social Knowledge Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E. and M.E. from Tohoku University, Miyagi in 1997 and 1999. Since joining NTT in 1999, he has been researching spoken dialogue systems, spoken language processing, and speech analytics. He joined NTT Communications in 2008 and was involved in the development of IP phone services on smartphones. He is a member of the Acoustical Society of Japan and JSAI.



Kazuyuki Iso

Senior Research Engineer, Social Knowledge Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E. from Gunma National College of Technology in 1998, and an M.S. and Ph.D. in knowledge science from Japan Advanced Institute of Science and Technology, Ishikawa, in 2000 and 2020. He joined NTT Cyber Space Laboratories in 2000. He is currently a senior research engineer at NTT Media intelligence Laboratories. His research interests include communication systems and computer-supported cooperative work. He is a member of the Institute of Electrical and Electronics Engineers (IEEE), the Virtual Reality Society of Japan (VRSJ), IPSJ, and IEICE.



Shuichi Nishioka

Senior Research Engineer, Supervisor, Social Knowledge Processing Laboratory, NTT Media Intelligence Laboratories.

He received a B.E. and Ph.D. in engineering from Yokohama National University, Kanagawa, in 1995 and 2005. He joined NTT in 1995. His current interests include data engineering and AI. He is a member of IPSJ and DBSJ.
