Technology Reports (Special Articles)

AI | Automatic Summarization | Natural Language Processing Techniques

Special Articles on AI—Expansion of AI Technologies to Diverse Industries and Basic Technologies Supporting AI Applications—

Multifunctional Automatic News Article Summarization AI System for Efficient Summarization

DOCOMO Beijing Communications Laboratories Xuhong Guo Issei Nakamura
Anxin Li
Service Innovation Department Hiroshi Fujimoto

Nowadays, many media outlets deliver summaries of news, which helps readers to understand news efficiently. However, manually writing a summary of each news article requires specialized skills, and leads to the issue of securing the human resources. NTT DOCOMO has developed a multifunctional AI system for automatic news article summarization. This system generates summaries based on the user's intention, reducing the time required for summarization and addressing the shortage of human resources.

1. Introduction

Nowadays, many media outlets provide simple summaries of news articles to present the main content to readers in an easy-to-understand manner. However, summaries often have a limit on the number of sentences or characters due to the area constraint on media outlets, and manually condensing news article into summaries with appropriate length requires specialized skills, in addition to handling large amounts of workloads. Therefore, it is necessary to train staff for a certain period of time, which is an issue from the standpoint of securing human resources.

To address this issue, in recent years a variety of automatic summarization systems using AI technology, which can improve the efficiency of summarization, have been released. Many of these systems use extractive summarization or abstractive summarization techniques. Extractive summarization extracts key sentences from the source text, whereas abstractive summarization creates new sentences by deleting or adding some words or phrases from the source text. Examples of extractive summarization and abstractive summarization are shown in **Figure 1**. The examples in all figures are provided only in Japanese since our system supports only Japanese at present.

Although having potential to generate high-quality news summaries in a similar way as a human, abstractive summarization has the following issues. Since some words or phrases are deleted or added to meet the required length of the summary, conventional technology often generates the summary with grammar errors. The summary may also not satisfy length requirements even if sentences are grammatically correct. In addition, many existing summarization systems do not have a function that takes into account keywords that the user wants to include or exclude, nor a function that visualizes the area of the source text from which the summary is copied or the area in the summary that is newly generated. As a result, with existing summarization systems users cannot add constraints to the summary content or efficiently correct grammatical errors in the summary, making them inconvenient to use in actual operation.

To solve the above issues, NTT DOCOMO has developed a multifunctional automatic summarization AI system that creates summaries in accordance with the user's intention. This system has three features. First, the character length control function creates summaries with a length of approximately 70 to 100% of the number of characters specified by the user. Compared with conventional methods, the precision of character length control and the quality of grammar of the summary have been improved. Second, our system has a hint

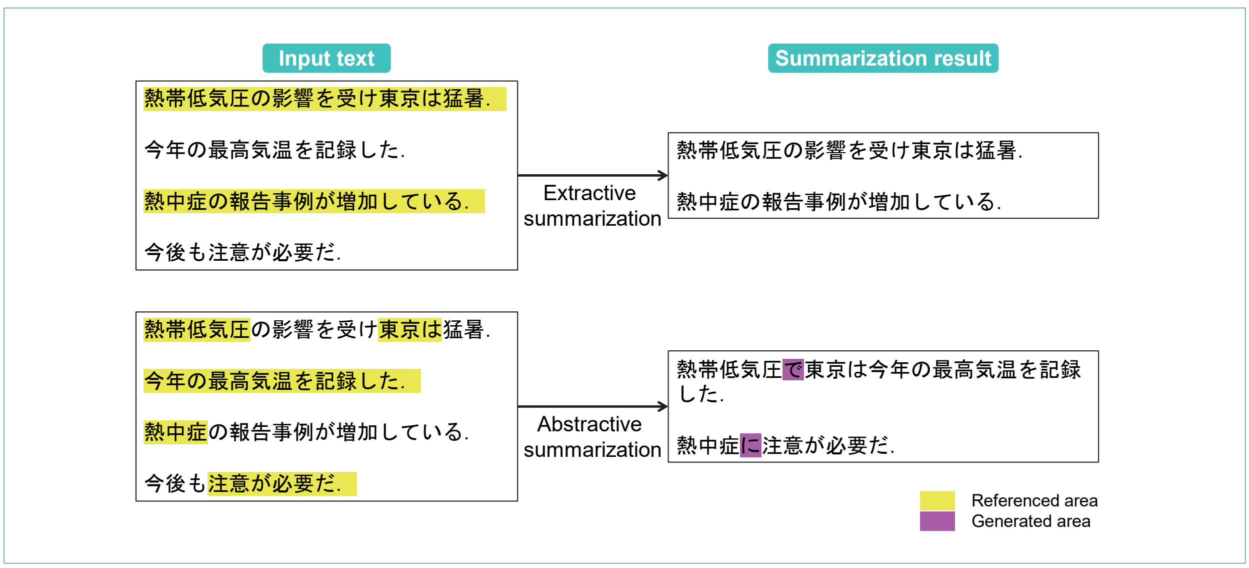

Figure 1　Examples of extractive summarization and abstractive summarization

function that allows users to set keywords to be included or excluded in the summary in order to efficiently generate summaries that meets the user's purpose. The system also has a title function that generates a summary that includes the content of the news article's title. Third, the system has the visualization function that highlights the areas in the source text that are referenced to generate summarization and the locations of newly generated text, which helps users to correct the summary more efficiently.

This system can automatically generate optimal summaries in terms of the number of characters and the content and allow the user to efficiently revise the generated summary when the summary includes errors. Compared with manual summarization and conventional summarization systems, this system can reduce the time required for summarization and addresses the shortage of human resources.

In this article, we describe the prominent functions of our automatic summarization AI system, techniques for improving the system's abstractive summarization performance, and the results of evaluation.

2. NTT DOCOMO's Automatic Summarization AI System

2.1 Overview of System

NTT DOCOMO's automatic summarization AI system is composed of two systems: an extractive summarization system and an abstractive summarization system (**Figure 2**). Both systems use deep

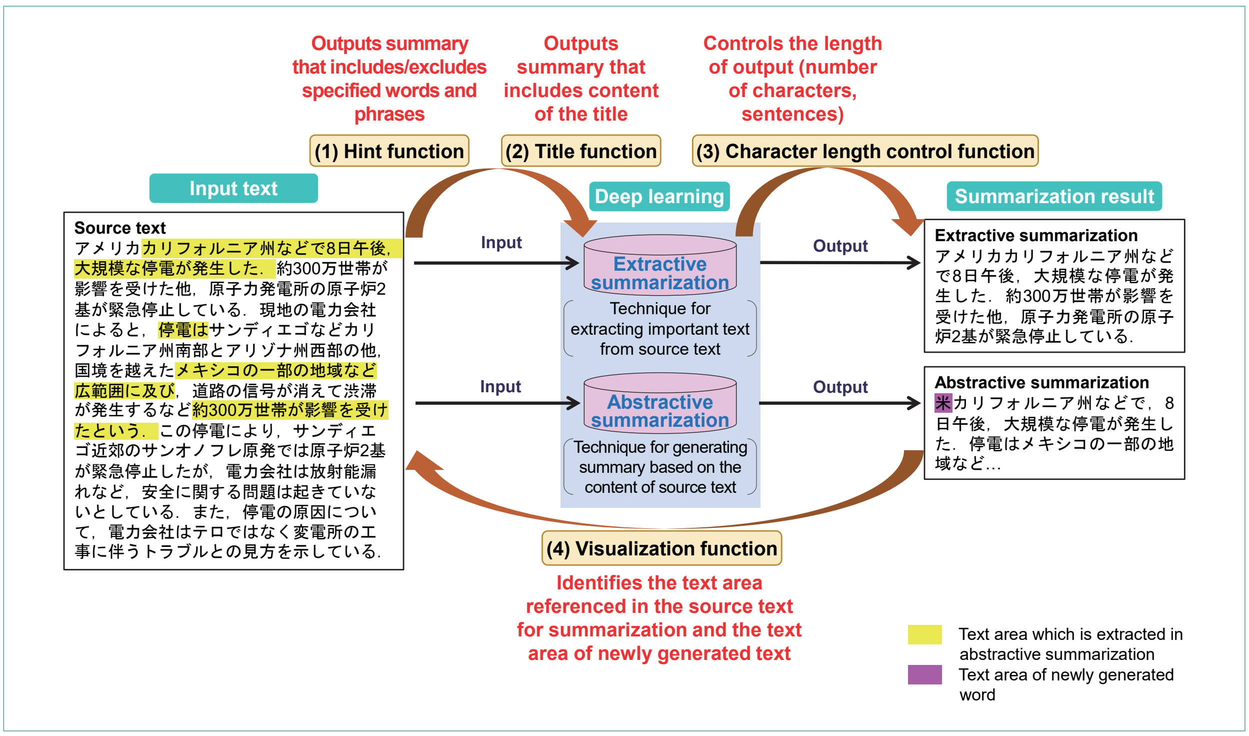

Figure 2 Overview of NTT DOCOMO's automatic summarization AI system

learning[*1] and are equipped with the following functions.

(1) Hint function: Adds restrictive conditions to the content of a summary through keywords and phrases specified by the user
(2) Title function: Generates summary that includes the content of the title
(3) Character length control function: Controls the number of characters and sentences in a summary
(4) Visualization function: Allows the user to visually check source text that were referenced to generate the summary and positions of newly generated text.

2.2 Hint Function

The hint function allows users to add restrictions to the content of a summary by specifying keywords or phrases they want to include or exclude. For a news article, the information that different users want to include in the summary, such as names of persons, places, and company appearing in the article, is different. This function was developed to generate summaries optimized for each user.

The system allows users to input multiple keywords and phrases. **Figure 3** shows an example of how the summary changes based on the keyword that the user wants to include in the summary. **Figure 4** shows an example of how the summary changes when the user wants to exclude the keyword suggested by the hint from the summary. In this way, the system makes it possible to output summaries that are closer to what users intend through input of hints.

2.3 Title Function

The title function can improve the quality of the summary by allowing the user to input the title of source text. The quality of deep learning-based summarization usually depends on the quality and quantity of the training data. When the source text contains content that is not included in the training data, the quality of the summary may decline. We thus developed the title function. This

Source text	Specified character length
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．約300万世帯が影響を受けた他，原子力発電所の原子炉2基が緊急停止している．現地の電力会社によると，停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他，国境を越えたメキシコの一部の地域など広範囲に及び，道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという．この停電により，サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが，電力会社は放射能漏れなど，安全に関する問題は起きていないとしている．また，停電の原因について，電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している．(294 characters)	120

Result of summarization without the keyword suggested by the hint
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．停電はメキシコの一部の地域など広範囲に及び，約300万世帯が影響を受けたという．また，停電の原因について，電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している．(118 characters)
Result of summarization including keyword“緊急停止”suggested by the hint
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．この停電により，サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが，電力会社は放射能漏れなど，安全に関する問題は起きていないとしている．(107 characters)

Keyword hint

Figure 3　Example of using hint function to produce abstractive summarization that includes hint

*1　Deep learning: A type of machine learning that uses multi-layer neural networks.

function improves the quality of summaries by using the title as a hint, which extracts important content from the source text. As shown in **Figure 5**, when the user specifies the title, the system generates a summary that includes keywords in the title.

2.4 Character Length Control Function

The character length control function allows the user to specify the number of characters in a summary. Because the size of a summary display area in a news website or social media is limited,

Source text	Specified character length
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．約300万世帯が影響を受けた他，原子力発電所の原子炉2基が緊急停止している．現地の電力会社によると，停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他，国境を越えたメキシコの一部の地域など広範囲に及び，道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという．この停電により，サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが，電力会社は放射能漏れなど，安全に関する問題は起きていないとしている．また，停電の原因について，電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している．(294 characters)	100

Result of summarization without excluding the keyword suggested by the hint
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．停電はメキシコの一部の地域など広範囲に及び，道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという．(90 characters)

Result of summarization excluding keyword “渋滞” suggested by the hint
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．この停電により，サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが，電力会社は，安全に関する問題は起きていないとしている．(100 characters)

Excluding keyword hint

Figure 4　Example of using hint function to produce abstractive summarization that excludes hint

Input	Specified character length
【タイトル】 米国などで大規模停電，原子炉2基緊急停止 【原文】 アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．約300万世帯が影響を受けた他，原子力発電所の原子炉2基が緊急停止している．現地の電力会社によると，停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他，国境を越えたメキシコの一部の地域など広範囲に及び，道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという．この停電により，サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが，電力会社は放射能漏れなど，安全に関する問題は起きていないとしている．また，停電の原因について，電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している．(294 characters)	100

Result of summarization without specification of title
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．停電はメキシコの一部の地域など広範囲に及び，道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという．(90 characters)

Result of summarization with specification of title
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．この停電により，サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが，電力会社は，安全に関する問題は起きていないとしている．(100 characters)

Title-related keywords

Figure 5　Example of abstractive summarization that uses title function there is a maximum number of the characters in the summary. This system thus is tuned so that the summary is kept within 70 – 100% of the character length specified by the user. As shown in **Figure 6**, when different character lengths are set for the same source text, the system generates summaries that satisfy the length requirement specified by the user. Using this function allows the system to generate summaries of various lengths. Besides the number of characters, the user can also specify the number of sentences.

2.5 Visualization Function

The visualization function allows the user to view the areas of the source text that are referenced to generate a summary and the positions of newly generated text in the summary. Existing automatic summarization AI systems do not show the areas of the source text that are referenced in the summary. As a result, it is time-consuming for the user to compare the source text and the summary to confirm whether or not the summary includes important content from the source text and whether or not the grammar of the generated sentences is correct. We therefore developed the visualization function, which highlights the areas of the source text referenced by the system to generate the summary and the positions of newly generated text. This function thus allows the user to efficiently compare the source text and summary result. As shown in **Figure 7**, the generated summary correctly mapped the phrase "原子炉2基が緊急停止" to the areas in the source text where it occurred even several times. Newly generated text is also correctly visualized.

Source text	Specified character length	Summarization result
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．約300万世帯が影響を受けた他，原子力発電所の原子炉2基が緊急停止している．現地の電力会社によると，停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他，国境を越えたメキシコの一部の地域など広範囲に及び，道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという．この停電により，サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが，電力会社は放射能漏れなど，安全に関する問題は起きていないとしている．また，停電の原因について，電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している．(294 characters)	80	アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．停電はメキシコの一部の地域など広範囲に及び，約300万世帯が影響を受けたという．(70 characters)
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．約300万世帯が影響を受けた他，原子力発電所の原子炉2基が緊急停止している．現地の電力会社によると，停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他，国境を越えたメキシコの一部の地域など広範囲に及び，道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという．この停電により，サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが，電力会社は放射能漏れなど，安全に関する問題は起きていないとしている．また，停電の原因について，電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している．(294 characters)	160	アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．停電はメキシコの一部の地域など広範囲に及び，道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという．また，停電の原因について，電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している．(136 characters)

Extracted position

Figure 6 Example of abstractive summarization that uses character length control function

Source text	Summarization result
アメリカカリフォルニア州などで8日午後，大規模な停電が発生した．約300万世帯が影響を受けた他，原子力発電所の原子炉2基が緊急停止している．現地の電力会社によると，停電はサンディエゴなどカリフォルニア州南部とアリゾナ州西部の他，国境を越えたメキシコの一部の地域など広範囲に及び，道路の信号が消えて渋滞が発生するなど約300万世帯が影響を受けたという．この停電により，サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが，電力会社は放射能漏れなど，安全に関する問題は起きていないとしている．また，停電の原因について，電力会社はテロではなく変電所の工事に伴うトラブルとの見方を示している．	米カリフォルニア州などで8日午後，大規模な停電が発生した．この停電により，サンディエゴ近郊のサンオノフレ原発では原子炉2基が緊急停止したが，電力会社は放射能漏れなど，安全に関する問題は起きていないとしている．

Area of source text referenced by abstractive summarization

Position of newly generated text

Figure 7　Example of visualization of text areas referenced for summarization and position of newly generated text

3. Methods for Improving Quality of Abstractive Summarization

3.1 Overview

Normally, the training of a deep learning-based summarization model uses a large amount of pair data of source text and human-created summaries. However, manually making summaries is time-consuming and costly.

Instead of requiring a large amount of prepared data, NTT DOCOMO's automatic summarization AI system improves the performance of the summarization model by using summaries containing errors generated mechanically from pair data consisting of source text and summaries as well as large amounts of source text data without correct summaries. It also uses compressed sentences from which extraneous information has been removed. Specifically, we introduced original technologies in the areas of grammar, non-redundancy, fluency, and character length control, as shown in **Figure 8** below. As a result, the system achieved greater performance compared with conventional methods.

(1) We generate sentences with grammatical errors mechanically and use reinforcement learning to reduce grammatical errors.

(2) We generate summaries with redundant content mechanically and deploy contrastive learning*2 to reduce redundancy.

(3) We implement pre-training to improve fluency by using a large amount of source text data.

(4) We deploy a sentence compression model trained using compressed sentences to control character length.

3.2 Grammar

Grammatical mistakes occur especially when summarizing patterns of word combinations that are not included in the training dataset. The cause is the use of "teacher forcing" algorithm [1], a technique widely used when training a deep learning model to generate text. In text generation using deep learning, the last generated word is used to

*2 Contrastive learning: A machine learning technique that increases the accuracy of the model by training the model to learn that the distances of features in similar data are closer than the distances between features in dissimilar data.

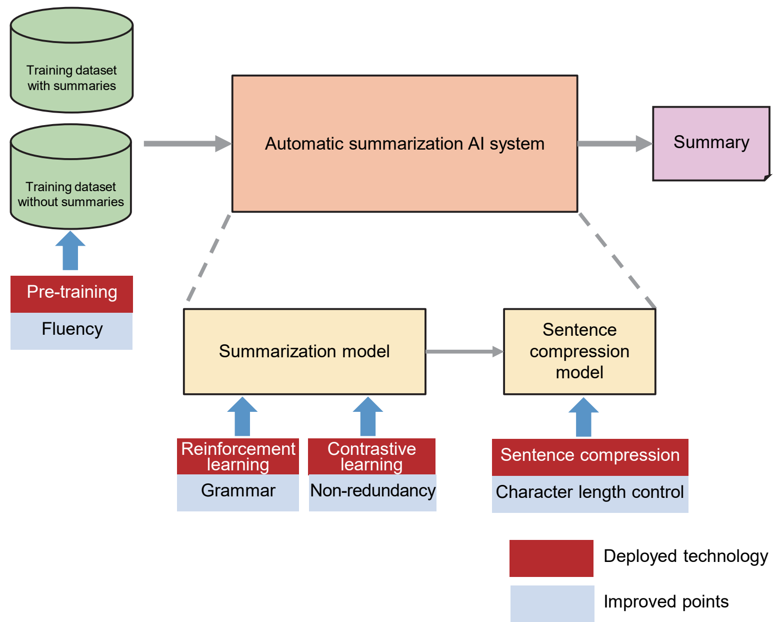

Figure 8　Summarization system's improved areas

generate the next word. Many incorrect words are generated in the initial stage of training, and teacher forcing can make training more efficient by directly inputting correct words into the deep learning decoder*3 instead of using the words previously generated by the model. In such training, because the model learns the next word to be generated for a specific word, word-level rather than entire sentence-level training is performed.

The issue is that even though the correct word is given during training, when generating a text using an unknown sentence, the text is generated using the word generated previously output by the model instead of the correct word. Therefore, once an erroneous word is generated, the model generates text that includes semantic and grammar errors.

To address this issue, NTT DOCOMO's automatic summarization AI system incorporates reinforcement learning. Reinforcement learning trains an intelligent agent through feedback between the agent and its environment, which provides the agent with rewards. While this technique is generally used in fields such as robotic control, in recent years it has also been used in the field of natural language processing [2]. In our system, reinforcement learning is composed of the following elements:

- Agent: The summarization model
- Environment: Discriminator that recognizes whether a given sentence is grammatically correct
- State: Summarization result
- Action: Generation of next word
- Reward: Score of grammatical correctness

*3 Decoder: In a multilayer neural network that takes input text and outputs new text, the part that receives text as input and converts it to feature values is called the Encoder; the part that generates new text by using the feature values output by the Encoder is called the Decoder.

(recognition result of by the discriminator)

By training the summarization model so that it gains the most reward through the action of generating summaries, in other words, by training the model to generate as few grammatical errors as possible, we developed a summarization model that produces fewer grammatical errors compared with conventional methods.

3.3 Non-redundancy

In this article, "redundancy" means that semantically identical content is repeated within the same sentence or across multiple sentences. Deep learning-based text generation models have been observed to have the issue of generating summaries that include redundant content [3].

NTT DOCOMO's automatic summarization AI system thus uses contrastive learning to address this problem and improves text generation performance. Contrastive learning is a technique for improving the accuracy of a model by training it with three datasets: an anchor dataset, positive examples, and negative examples. During training, the model learns features, considering distances between the feature values of the anchor dataset and the positive examples as closer than the distances between the feature values of the anchor dataset and negative examples. In the NTT DOCOMO's automatic summarization AI system, anchors are the summaries generated by the summarization model. Positive examples are correct human-created summaries, and negative examples are incorrect summaries in terms of redundancy, with the same words, phrases, and sentences repeated mechanically. By training the model using negative, erroneous examples in terms of redundancy, we developed a summarization model that generates fewer redundant sentences compared with conventional summarization methods.

3.4 Fluency

In this article, "fluency" expresses the quality of combinations of words and phrases being correct. In actual use, our system may receive as input source text that contains combinations of words that are not included in the training dataset. In such a case, conventional methods may generate a summary with a combination of words that contains wrong usage while being grammatically correct.

To address this issue, NTT DOCOMO's automatic summarization AI system conducts pre-training on word combinations using a large amount of text, resulting in improved fluency in the summarization model.

3.5 Character Length Control

The most common technique for character length control is to choose the output result with the appropriate length based on rules when conducting a beam search*4 in the process of generating a summary. In such a rule-based method, grammatical correctness and the degree of pertinence are not fully considered when selecting an output with an appropriate length. It is thus easy to output a summary that, while close to the length specified by the user, contains grammatical errors or strays from the main subject of the source text [4].

To address this issue, NTT DOCOMO's character length control function inputs the summary length information to the model as a feature value*5 in order to optimize the model to simultaneously

*4 Beam search: In this article, refers to the selection of multiple candidate words output by the neural network based on their scores, resulting in several summarization result candidates.

*5 Feature values: Values extracted from data, and given to that data to give their features.

learn grammar, pertinence, and length. In addition, as a post-processing technique, a text compression model is applied to reduce the number of characters in the summary. If the output exceeds the character length specified by the user, the system can compress the summary to the appropriate length.

4. Evaluation of Performance of Abstractive Summarization

4.1 Dataset

We used a dataset of about 180,000 news articles and their human-created summaries, provided by the Nippon Television Network Corporation, as the dataset for training and evaluating our system's abstractive summarization model.

4.2 ROUGE Evaluation

Figure 9 shows how each metric of Recall-Oriented Understudy for Gisting Evaluation (ROUGE) is calculated. ROUGE evaluates the comprehensiveness (recall) of a summarization by comparing with a correct summary (reference). ROUGE is the most widely used metric for evaluating summarization

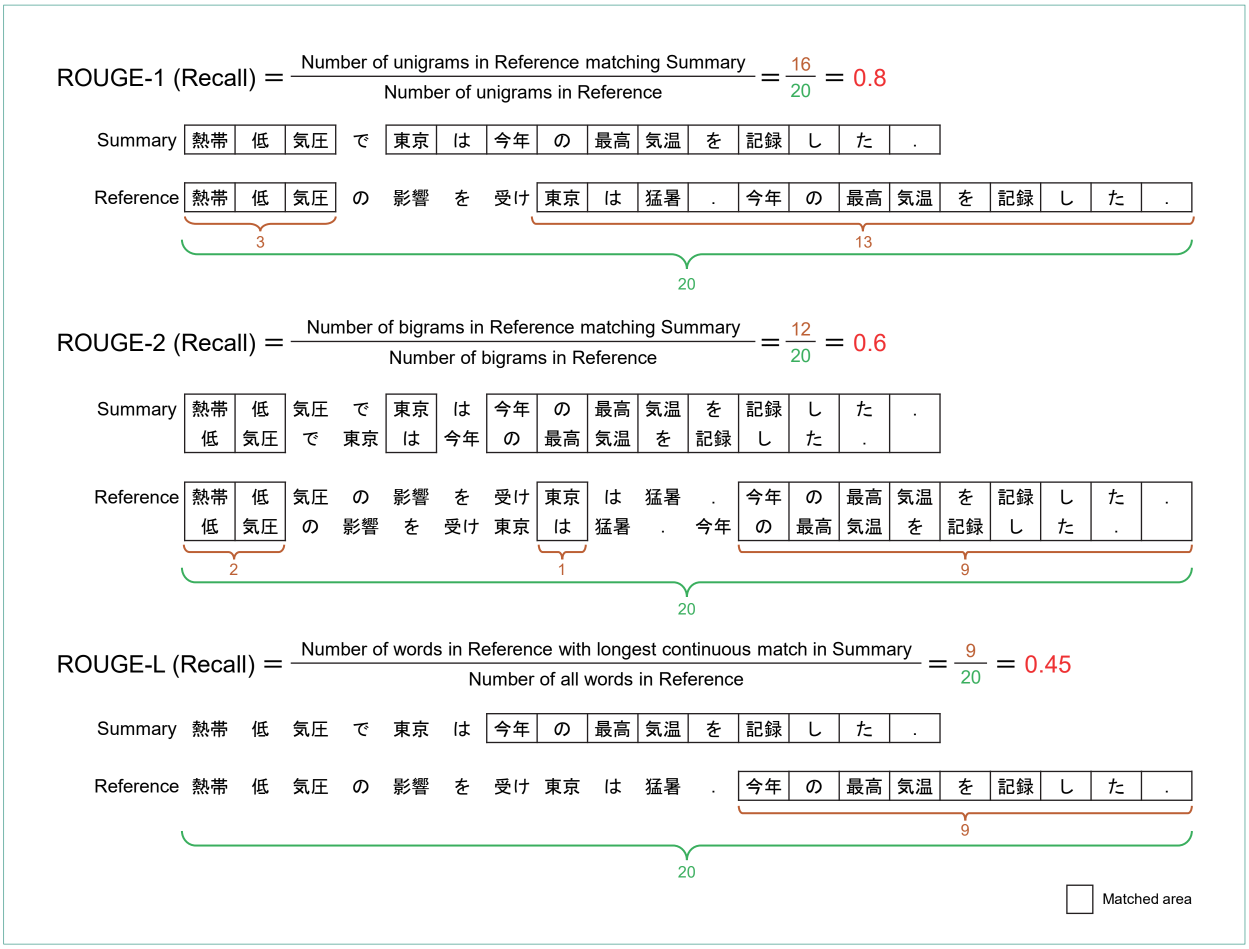

Figure 9 ROUGE metrics calculation methods

by comparing the similarity of text generated by a model with the reference text [5]. ROUGE-1 and ROUGE-2 represent the degree of overlap of unigrams*6 and bigrams*7, respectively, between the generated text and the reference text. ROUGE-L measures the degree of overlap using the length of longest common text fragment. For all ROUGE metrics such as ROUGE-1, ROUGE-2, and ROUGE-L, the higher the score, the greater the degree of overlap between the two texts, and the greater the performance of the text generation model.

Table 1 shows the evaluation results of NTT DOCOMO's summarization system using 3,000 texts from Nippon Television Network Corporation's dataset. We compared our summarization model with Lead-3, which takes the first three sentences of the source text as the summarization result, as well as TextRank [6], SumBasic [7], LSA (Latent Semantic Analysis) [8], Submodular [9], and PGN (Pointer Generator Network) [3]. TextRank, SumBasic, LSA, and Submodular are extractive summarization techniques and extract the three most important sentences in the source text as the summarization result. PGN is an abstractive summarization technique. The summary closest to the character length of the most correct summary in beam search results is evaluated as the final summarization result. Compared to these summarization methods, NTT DOCOMO's automatic summarization AI system achieved higher ROUGE scores, indicating that it generates summaries close to correct, human-created summaries.

4.3 Human Evaluation

One of the disadvantages of ROUGE evaluation described above is that it cannot evaluate grammatical mistakes and semantic redundancies. Thus, a person whose native language is Japanese evaluated summaries using the following four criteria. Each criterion was evaluated on a 4-point scale, with 4 being the highest score.

(1) Grammar: Few grammatical mistakes in the summary

(2) Pertinence: The generated summary covers the main content of the source text

(3) Non-redundancy: Semantically identical words, phrases, and sentences are not repeated in the summary

Table 1 Results of ROUGE evaluation

		ROUGE-1	ROUGE-2	ROUGE-L
Extractive summarization	Lead-3	74.46	63.89	72.48
	TextRank	64.06	50.07	60.16
	SumBasic	64.49	49.18	58.38
	LSA	62.28	46.48	56.85
	Submodular	55.41	36.91	47.41
Abstractive summarization	PGN	79.25	70.36	77.45
	NTT DOCOMO	84.49	76.47	81.80

*6 Unigram: A string of n consecutive words is called an n-gram. When n is 1, the string consists of only one word.

*7 Bigram: A string of two consecutive words.

(4) Fluency: The flow from word to word and sentence to sentence is smooth in the generated summary

Table 2 shows the average scores of human evaluation of 100 summaries produced from the evaluation dataset by NTT DOCOMO's automatic summarization AI system and PGN. An additional score on a 4-point scale is calculated for each summary's character length; the criteria for each score, based on attainment of the specified character length, are given in **Table 3**. This score is calculated for each generated summary, and the average of all the scores is added to the last column "Character length" in Table 2. The results of Table 2 show that NTT DOCOMO's automatic summarization AI system scores higher than the PGN method in all evaluation metrics, indicating that it generates higher-quality summaries. Note that our system's automatic summarization takes about 1 second for extractive summarization and 10 seconds for abstractive summarization. These performances are much faster than human summarization, which requires several minutes.

5. Conclusion

In this article, we described the functions, techniques, and performance of NTT DOCOMO's automatic summarization AI system. NTT DOCOMO has developed the following functions to realize a system that easily outputs summaries according to the user's intentions: a hint function, which allows the user to specify hints when summarizing;

Table 2　Results of human evaluation

	Grammar	Subject matter pertinence	Non-redundancy	Fluency	Number of characters of summary
PGN	2.82	2.40	3.89	3.06	3.25
NTT DOCOMO	3.85	3.53	3.92	3.84	3.89

Table 3　Score related to number of characters of generated summary

Range of the number of characters in summary	Score
$0.7 \times L \leqq S \leqq 1.0 \times L$	4
$0.6 \times L \leqq S < 0.7 \times L$ or $1.0 \times L < S \leqq 1.1 \times L$	3
$0.5 \times L \leqq S < 0.6 \times L$ or $1.1 \times L < S \leqq 1.2 \times L$	2
$S < 0.5 \times L$ or $1.2 \times L < S$	1

L: The number of characters specified by user

S: Character count of summary generated by the summarization model a title function, which allows the title to be used in summarizing; a character length control function, which controls the number of characters and sentences in a summary; and a visualization function, which helps the user to view the areas of text referenced in the source text and the positions of newly generated text. To solve issues related to grammar, non-redundancy, fluency, and summary character length control, we deployed techniques such as reinforcement learning, contrastive learning, and pre-training; these techniques led to improvement of performance. As the results of performance evaluation show, our automatic summarization AI system greatly outperformed conventional methods in ROUGE metrics and human evaluation. NTT DOCOMO's automatic summarization AI system can reduce the time required for summarization and address the shortage of human resources. We will continue to improve the performance of existing functions and develop new functions to respond to issues experienced in actual service and realize an even higher-performing automatic summarization AI system.

REFERENCES

[1] R. J. Williams and D. Zipser: "A Learning Algorithm for Continually Running Fully Recurrent Neural Networks," Neural Comput., Vol.1, No.2, pp.270–280, 1989 (doi: 10.1162/neco.1989.1.2.270).

[2] L. Yu, W. Zhang, J. Wang and Y. Yu: "SeqGAN: Sequence generative adversarial nets with policy gradient," 31st AAAI Conf. Artif. on Intell., pp.2852–2858, 2017.

[3] A. See, P. J. Liu and C. D. Manning: "Get To The Point: Summarization with Pointer-Generator Networks," ACL 2017 - 55th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap.), Vol.1, pp.1073–1083, 2017 (doi: 10.18653/v1/P17-1099).

[4] B. Eikema and W. Aziz: "Is MAP Decoding All You Need? The Inadequacy of the Mode in Neural Machine Translation," Proc. of the 28th International Conference on Computational Linguistics, pp.4506–4520, 2021 (doi: 10.18653/v1/2020.coling-main.398).

[5] C.-Y. Lin: "Looking for a Few Good Metrics: ROUGE and its Evaluation," NTCIR Work., pp.1–8, Jun. 2004.

[6] R. Mihalcea and P. Tarau: "TextRank: Bringing Order into Text," Proc. of 2004 Conf. Empir. Methods Nat. Lang. Process., pp.404–411, 2004.

[7] L. Vanderwende, H. Suzuki, C. Brockett and A. Nenkova: "Beyond SumBasic: Task-focused summarization with sentence simplification and lexical expansion," Inf. Process. Manag., Vol.43, No.6, pp.1606–1618, Nov. 2007 (doi: 10.1016/j.ipm.2007.01.023).

[8] M. G. Ozsoy, F. N. Alpaslan and I. Cicekli: "Text summarization using latent semantic analysis," J. Inf. Sci., Vol.37, No.4, pp.405–417, Aug. 2011 (doi: 10.1177/0165551511408848).

[9] H. Lin and J. Bilmes: "A class of submodular functions for document summarization," Proc. of 49th Annu. Meet. Assoc. Comput. Linguist. Hum. Lang. Technol., Vol.1, pp.510–520, Jun. 2011.