URL　https://journal.ntt.co.jp/article/22763
DOI　https://doi.org/10.60249/23085102

要約作業を効率化する多機能なニュース記事自動要約AIシステム

現在，多くのメディアのニュースサイトでは，記事の要約を表示することにより読者に内容を分かりやすく伝えています．しかし，人手による要約作業は分量が多いうえに専門的なスキルを必要とし，人材確保の観点で課題となっています．そこでNTTドコモは，要約作業の効率化をめざし，ユーザの意向に沿った要約を自動で作成し，さらに確認作業を効率化できる多機能な自動要約AI（人工知能）システムを開発しました．特に，汎用的な対話型AIでは現状は実現が困難な位置特定機能を活用することで，より効率的に要約を作成できます．

キーワード：#AI，#自動要約，#自然言語処理技術

郭 塽宏[†1] / 中村 一成[†1]
李 安新[†1] / 藤本 拓[†2]

NTTドコモ北京研究所[†1]
NTTドコモ[†2]

はじめに

人手による要約作業の効率化をめざし，NTTドコモはユーザの意向に沿った要約を作成する多機能な自動要約AI（人工知能）システムを開発しました．本システムはヒント機能，文字数制御機能，位置特定機能を具備し，文字数や内容の観点でユーザの求める最適な要約を自動で生成することが可能です．本システムを活用することで，効率的に要約の自動生成，確認と修正作業ができるため，人手による要約作成や従来の要約システムを利用した場合と比較して要約作業に要する時間を短縮することが可能となります．

NTTドコモの自動要約AIシステム

■システム全体像

NTTドコモの自動要約AIシステムには，深層学習[*1]を用いた抽出式要約と生成式要約の2つのシステムがあります（**図1**）．各システムに実装されている機能を以下で解説します．

■ヒント機能

ヒント機能は，ユーザが要約に含めたい，もしくは含めたくないキーワードやフレーズを指定することで，要約内容に制約を加えることができる機能です．ニュース記事において，どの情報を要約に含めたいかはユーザごとに異なる可能性があるため，各ユーザに対して最適な要約を生成できるよう，このような機能を開発しました．

図2 (a)は要約に含めたいキーワードをユーザが指定することで，ヒントがない場合には要約に含まれない内容が含まれるように変化させた例です．図2 (b)は要約に含めたくないキーワード（図中ではマイナス

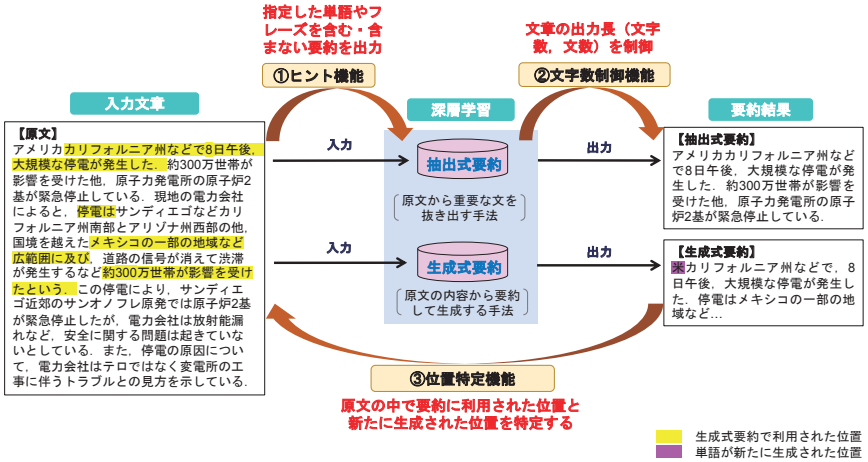


図1　自動要約AIシステム全体像

(a)

原　文	指定文字数
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文字）	120

ヒント指定なし　要約結果
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文字）

ヒント指定“緊急停止”　要約結果
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文字）

(b)

原　文	指定文字数
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文字）	100

マイナスヒント指定なし　要約結果
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文字）

マイナスヒント指定“渋滞”　要約結果
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文字）

ヒント　　マイナスヒント

図2　ヒント機能

＊　本記事は「NTT DOCOMOテクニカル・ジャーナル」（Vol.29 No.4, 2022年1月）に掲載された内容を編集したものです．

＊1　深層学習：多層のニューラルネットワークを用いた機械学習の一種．

原　文	指定文字数	要約結果
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文字）	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文字）
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文字）	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文字）

引用部分

図3　文字数制御機能

原　文	要約結果
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．

生成式要約で引用された位置
単語が新たに生成された位置

図4　位置特定機能

ヒントと記載）をユーザが指定することで，ヒントがない場合に要約に含まれていた内容が，含まれないように変化した例です．

■文字数制御機能

文字数制御機能は，ユーザが要約の文字数を指定できる機能です．ニュースサイトやソーシャルメディアには，配信メディアの要約表示枠の制約から要約文字数の最大値が厳しく制限されることがあるため，本システムではユーザが設定した文字数の70～100％の範囲で要約を出力するようにチューニングをしています．**図3**に示すように同じ原文に対して異なる文字数を設定すると，それに合った文字数の要約をそれぞれ得ることができます．

■位置特定機能

位置特定機能は，要約の文言が原文のどの位置を参照したかを可視化できる機能です．要約が原文の中の重要な内容を含み，文法的に正しいかどうかをユーザが効率的に確認できるようにこのような機能を開発しました．**図4**に示すように，「原子炉2基が緊急停止」という同じフレーズが原文に複数ある場合や，単語が新たに生成された場合にも正しく参照個所をマッピングすることが可能です．このような各文言のマッピングは汎用的な対話型AIでは現状は実現が困難であるため，本システムを活用することで，汎用的な対話型AIを活用した要約作業と比較してさらに効率的に要約を作成できます．

生成式要約の品質改善手法

■概　要

深層学習を用いた要約モデルの学習では，通常，原文と人手で作成した正しい要約のペアのデータを用いて学習を行いますが，人手で要約を作成するのは時間的，また金銭的なコストがかかることが課題です．

そこで，大量のデータを用意する代わりに，原文と要約のペアデータから機械的に生成された誤りのある要約や，正解の要約がない大量の原文データ，一文の中から不要な情報が除かれた圧縮文を活用し，独自の技術を導入することで，既存技術と比較して主に文法，非冗長性，文字数制御の観点で，性能改善を行いました．

■文　法

本システムでは，強化学習の方法を導入することで文法の性能改善を行いました．強化学習はエージェントとそれに報酬を返す環境の間のフィードバックをとおして，エージェントを学習させる手法です．一般的にロボット制御などの分野で利用される学習手法ですが，近年自然言語処理の分野でも利用される例があります[(1)]．本システムでは，強化学習の各要素が下記のように構成されます．

- エージェント（Agent）：要約モデル
- 環境（Environment）：文を与えたときに文法的に正しいか正しくないかを判断する識別器
- 状態（State）：要約結果
- 動作（Action）：次の単語の生成
- 報酬（Reward）：文法の正しさのスコア（識別器の判断結果）

要約モデルが要約を生成する動作を通じて，報酬がもっとも多くなるように，つまり文法の誤りが少なくなるように学習を行うことで，既存手法と比較して文法誤りの少ない要約モデルを構築しました．

■非冗長性

本稿では要約内容の中で，同一文内や複数の文間で意味的に同一の内容が繰り返し言及されることを「冗長」と表現します．深層学習を用いたテキスト生成モデルでは，冗長な内容を含む要約の生成が課題として指摘されています[(2)]．

そこで，本システムでは，対照学習＊2を導入することにより冗長性の課題について改善を行いました．対照学習では，Anchor，正例，負例の3つを用意し，学習時にAnchorと正例の特徴間の距離が，Anchorと負例の特徴間の距離よりも近接するように学習することでモデルの精度を

＊2 対照学習：類似データ間の特徴距離が非類似データ間の特徴距離よりも近くなるように学習することでモデルの高精度化を行う手法．

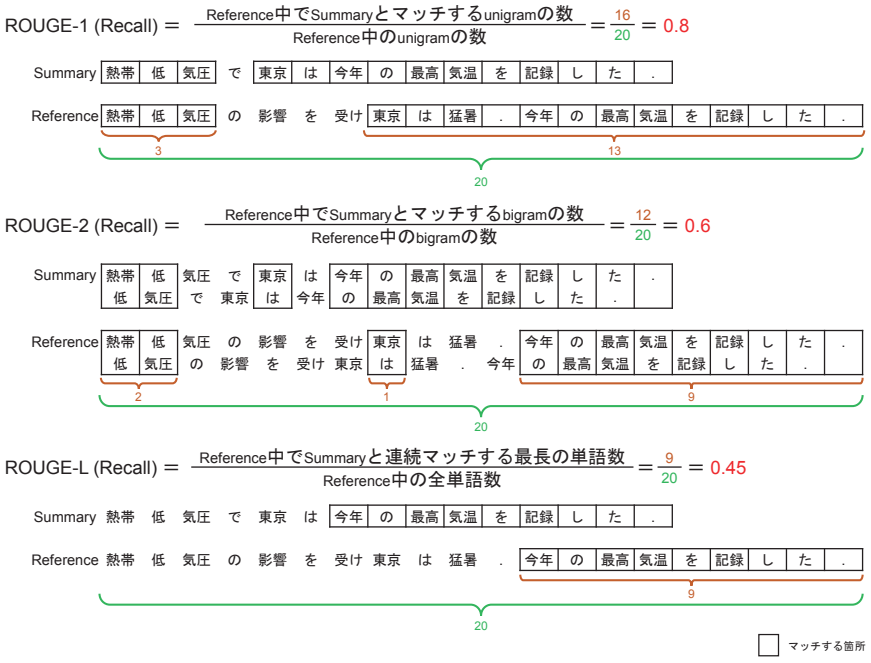


図 5　ROUGEの計算方法

表 1　ROUGEによる評価結果

		ROUGE-1	ROUGE-2	ROUGE-L
抽出式要約	Lead-3	74.46	63.89	72.48
	TextRank	64.06	50.07	60.16
	SumBasic	64.49	49.18	58.38
生成式要約	PGN	79.25	70.36	77.45
	Ours	84.49	76.47	81.80

向上する手法です．本システムでは，Anchorは要約モデルによって生成された要約，正例は人手で作成した正しい要約であり，負例は正例に対して機械的に同一の単語，フレーズ，文を繰り返した，冗長性の観点で誤りのある要約としました．このように冗長性の観点で誤りのある負例を用いた学習を行うことで，既存手法と比較して冗長な文を生成しにくい要約モデルを構築しました．

■文字数制御

文字数制御でもっとも一般的な手法は，要約を生成する過程でBeam Search＊3を行う際に，ルールベースでより適切な長さとなる出力結果を選択する手法です．このようなルールベースの手法の場合，適切な長さとなる出力結果を選択するために，文法的な正しさや主題の抽出度合いについて考慮されない場合があり，要約結果がユーザの指定した長さに近い一方で，文法的に誤りのある要約や主題から逸脱した要約が出力されやすくなる課題があります(3)．

そこでNTTドコモの文字数制御機能は，要約モデル自体に出力する要約の長さの情報を特徴量＊4として入力し，要約モデルが要約を出力する際に文法，主題，長さなどを同時に考慮するよう学習させることで，要約モデルの最適化を行いました．また要約モデルの後処理として，文字数を削減するための文圧縮モデルを導入することにより，要約モデルより出力された文がユーザの設定する文字数を上回る場合は，文を適切な長さに圧縮し要約を生成できるように改善しました．

生成式要約の性能評価

■データセット

日本テレビ放送網株式会社が提供する約18万件のニュース記事と記事を人手で要約したデータセットを利用し，本システムの生成式要約の学習と評価を行いました．

■ROUGEによる評価

正解要約（Reference）に対する要約結果の網羅性を評価するRecallについて各ROUGEの計算方法を**図 5** に示します．ROUGEは，正解とされるテキストとモデルが作成したテキストの類似度を比較する要約の指標の中でもっとも広く利用されるものです(4)．ROUGE-1，ROUGE-2は，それぞれテキスト間のunigram＊5とbigram＊6の重なり度合いを表し，ROUGE-Lは，一致する最長の単語の長さを用いてテキスト間の重なり度合いを測定します．ROUGE-1，ROUGE-2，ROUGE-Lすべてにおいて，値が大きいほどテキスト間の重なりが大きく，テキスト生成モデルの性能が高いことを示します．

表 1 は，日本テレビ放送網が提供するデータセットのうち3000件を用いて，本システムの評価を行った結果です．比較対象として原文の先頭3つの文を要約結果とするLead-3，また，TextRank(5)，SumBasic(6)，PGN(2)を利用しました．TextRank，SumBasicについては抽出式要約であり，抽出する文の中で重要度の高い3つの文を要約結果としました．PGNについては生成式要約であり，Beam Search結果の中でもっとも正解の要約の文字数に近い要約を，最終的な要約結果として評価しました．これらの既存手法と比較して，NTTドコモの自動要約AIシステムはROUGEの数値が大きく，人手で作成した

＊3 Beam Search：本稿では，ニューラルネットワークが出力する単語の候補をスコアに基づいて複数選定し，いくつかの要約結果の候補を得ること．
＊4 特徴量：データから抽出される，そのデータを特徴付ける量（数値）．
＊5 unigram：n単語連続して続く文字列をn-gramと呼び，nが1の場合の1単語だけの文字列のこと．
＊6 bigram：2単語連続して続く文字列のこと．

表 2　人手による評価結果

	文法	主題性	非冗長性	流暢性	要約文字数
PGN	2.82	2.40	3.89	3.06	3.25
Ours	3.85	3.53	3.92	3.84	3.89

表 3　要約の文字数に関するスコア

要約の文字数範囲	スコア
$0.7 \times L \leqq S \leqq 1.0 \times L$	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：正解要約の文字数
S：要約モデルが作成した要約の文字数

正解により近い要約を生成することが示されました．

■人手による評価

前述のROUGEの評価による欠点として，文法の誤りや意味的な冗長性などを評価できない点が挙げられます．そこで，日本語を母国語とする者が以下の 4 つの観点で評価を行いました．その際スコアを，4 を最高点とする 4 段階評価としました．

① 文法：要約に文法的な誤りが少ないこと
② 主題性：生成された要約が原文の主要な内容をカバーしていること
③ 非冗長性：要約の中に意味的に同一の単語，フレーズ，文の繰返しがないこと
④ 流暢性：生成された要約が単語間，文間で流暢であること

表 2 は，評価データセットのうち100件を用いてNTTドコモの自動要約AIシステムとPGNの出力結果を人手で評価した評価値の平均値です．また**表 3** に示すように，指定した要約の文字数に対して生成された要約の文字数の割合を算出し，4 段階のスコアに割り当てました．この要約文字数のスコアを評価データごとに算出し，平均値を表 2 の「要約文字数」列に記載しました．

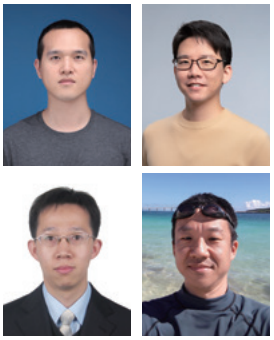
このような人手による評価についても既存手法のPGNと比較して，NTTドコモの自動要約AIシステムは各評価指標の数値が大きく，より品質の高い要約が生成されることが示されました．なお，自動要約の処理時間は，抽出式要約が約 1 秒，生成式要約が10秒程度のため，数分から数十分かかる手動の要約作業と比較して速度が大幅に向上します．

おわりに

本稿では，NTTドコモの自動要約AIシステムの各機能や技術，性能について解説しました．NTTドコモは，ヒント機能，文字数制御機能，位置特定機能を開発し，ユーザの意向に沿った要約を出力しやすいシステムを実現しました．また，文法，非冗長性，要約文字数の制約の課題を解決するために，強化学習，対照学習などを導入し，性能向上を行いました．性能評価結果が示すように，NTTドコモの自動要約AIシステムは既存の技術と比較し，ROUGEによる評価と人手による評価の両方の指標の値を改善しました．この自動要約AIシステムにより，要約作業時間を短縮し人手不足の解消をすることが可能となります．今後は，サービスをとおして得られた課題を基に既存機能の性能向上や新機能の開発を行い，さらに高性能な自動要約AIシステムを実現していきます．

■参考文献

(1) 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.
(2) 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).
(3) 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).
(4) C.-Y. Lin: "Looking for a Few Good Metrics: ROUGE and its Evaluation," NTCIR Work., pp.1–8, June 2004.
(5) R. Mihalcea and P. Tarau: "TextRank: Bringing Order into Text," Proc. of 2004 Conf. Empir. methods Nat. Lang. Process., pp. 404-411, 2004.
(6) 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).



（上段左から）郭　塽宏／中村　一成
（下段左から）李　安新／藤本　拓

今回紹介したシステムはパートナー企業との密な連携により，現場で要約を行っている方々のニーズに基づいて開発を行いました．今後もパートナー企業との連携を推進し，社会課題の解決に貢献していきます．

◆問い合わせ先
NTTドコモ
R&D戦略部
E-mail　dtj @ nttdocomo.com