2 文書要約技術

テキストの内容を的確に理解し、要点をおさえた自然な文章で提供する文書要約技術

NTT 人間情報研究所（以下、人間研）の共生知能研究プロジェクト（以下、共生 P）では、長文のテキストを読み解き、要約文を自動で生成する技術を研究開発している。AI の活用により、超大規模データから学習した事前知識を活用した重要語抽出モデルにより、入力テキストの重要な箇所を的確に判断するとともに、自然な文章の要約文を生成できる技術を紹介する。

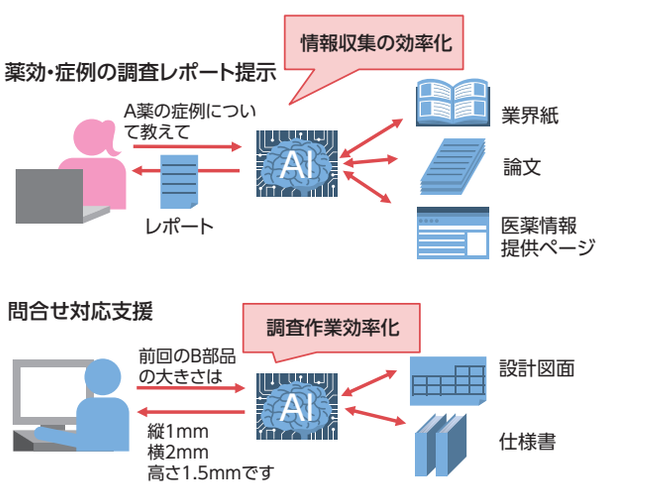
NTT 人間情報研究所
共生知能研究プロジェクト
（左）研究主任　**川場 真理子**氏
（右）主幹研究員　**齋藤 邦子**氏

人の機能を代替・補助する言語処理技術

人間研の共生 P では、大量・多様な文書を読み解く言語処理技術に取り組んでいる。これは従来人間が時間をかけて行う必要があった作業を代替・補助することをめざすもので、大量のマニュアルを読み解いて質問に対して回答を的確に探し出す機械読解技術、長いテキストの内容の要点をおさえて端的にまとめる要約技術などがある。これらの技術により、問い合わせ対応や情報収集の効率化が期待できる。

例えば、医薬関連の現場において予め AI に業界紙・論文・医療情報提供ページ等を学習させておくことにより薬能や症例の情報収集の効率化を図ったり、設計関連の現場において図面や仕様書を学習させておくことで調査作業の効率化を図ったり、といったことが可能になる（図 1）。つまり、これまで人間が時間をかけて行っていた情報のまとめ、判断、作業の一部を AI に任せることで、作業効率化や生産性の向上に貢献することをめざす。

以下に、用途に合わせて要点をおさえた自然な要約文を生成する自動要約技術について紹介する。



効果・価値
・大量・多様な文書を読み解く言語処理技術により、従来人間が時間をかけて実施している情報収集や調査作業の効率化を実現する

図 1　人の機能を代替・補助する言語処理技術の実現イメージと効果

出力長の制御が可能な自動要約技術

人間研で実施している文書要約技術の概要を示す（図 2）。

この技術では、出力したい要約の長さを指定すると、指定した長さに応じた個数の重要語を抽出し、その情報を元に要約を生成する。そのため、任意の要約長でも重要な情報を落とさず、精度良く要約を作成することができる。

スマートフォンで表示する場合には 50 文字程度の長さで、タブレットは表示可能面積が大きいので 150 文字程度の長さでというように、表示するデバイスに合わせた長さ調整

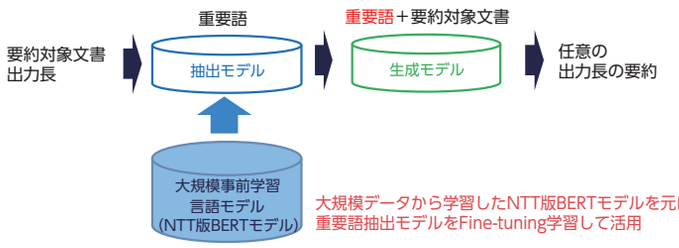


図２　サービス展開中のシステム概要

が可能である。

なお、この自動要約技術は、NTTコミュニケーションズ社へ技術提供しており、同社から「COTOHA Summarize[1]」という名前でサービス展開されている。

大規模汎用日本語言語モデル

人間研の要約技術を支えているものが大規模汎用日本語言語モデルである。近年、大量のテキストデータを事前学習する大規模汎用言語モデル（BERT[2] や MASS[3]）の研究が活発に行われている。これらの事前学習モデルに更に分類や抽出などの言語処理タスクに依存する教師あり学習（Fine-tuning）を追加実施することで、様々な言語処理タスクでの劇的な精度向上が報告されている。

NTT 人間研では、長年の日本語言語処理に関するノウハウを活かし、独自に大規模日本語テキストデータを準備し、NTT 版 BERT モデルを整備した。図２の要約技術では NTT 版 BERT モデルに基づく重要語抽出タスクの Fine-tuning を実施した抽出モデルにより、任意の出力長で要約に重要な語を精度よく抽出する機能を実現している。

今後の自動要約技術の研究開発にむけて

現在、人間研では、要約技術の改良に取り組んでいる。

１点目は、抽出モデルが予測する重要語に加えてユーザが指定する任意のキーワードを観点として考慮する機能の追加である。これにより、ユーザが着目したポイントに合わせた要約を生成できるようになる。

２点目は、要約の入力テキストの長文対応である。従来技術では、抽出モデルの元になる BERT モデルの制約で 700 文字程度の長さまでしか扱えなかったが、これを長文に対応できるようにし、6000 文字程度と大幅に伸ばすことができる。これにより、従来は長めの文書は事前に分割するなど前処理が必要であったが、前処理なく丸ごと要約処理を行うことが可能となる。

3点目は、生成モデルへの大規模事前学習言語モデル（NTT 版 MASS モデル）の採用である。これにより、要約生成の精度向上が期待できる。

これらの改良によって、より高精度かつ、多様な文書を要約でき、さらに要約結果にユーザの意図を反映させることができるようになることをめざしている。要約できるようになることをめざしている。

参考資料
[1] https://api.ce-cotoha.com/contents/index.html
[2] Devlin et al., BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, NAACL, 2019
[3] Song et al., MASS: Masked Sequence to Sequence Pre-training for Language Generation, ICML, 2019

研究の出口

2019 年に Google 社から「BERT」が提案され、以降いろいろな組織が日本語版の BERT を出すようになりました。人間研には言語資源が大量にあるので、その強みを生かして大規模な言語モデルを作りました。これまで研究所は日本語の解析に強みがあり、言語解析器など自社製のものを多く作ってきました。しかし、機械学習技術の発展によって言語の壁が低くなってくるとともに既存の OSS も良いものがたくさん出てくるようになったので、これまで培ってきた日本語の解析のノウハウ、人間研独自の技術を OSS にどのように組み込んでいくかを検討する必要があると考えています。利用できるものをうまく取り入れて、スピード感をもって人間研の技術の外部展開をめざしていきたいです。

現時点ではニュースや原稿などきちんとした文章を要約する場合が多いと思いますが、今後の要約の研究の目標としては、人と人の対話を要約できないかと考えています。例えば、オペレーターとお客さんがやり取りする局面などの、くだけた話し言葉や２者のやりとりを理解して要約できるようになると利用範囲が広がるのではないかと考えています。　（川場氏、齋藤氏　談）