

2020年1月21日
日本テレビ放送網株式会社
株式会社NTTドコモ

AIを活用したニュース記事の自動要約における実証実験を実施 ～実用化に向け、自動要約システムのプロトタイプを開発～

日本テレビ放送網株式会社(以下、日本テレビ)と、株式会社NTTドコモ(以下、ドコモ)は、AIを活用したニュース記事の自動要約システム(以下、本システム)の実用化をめざす実証実験(以下、本実証実験)を行い、WEB用ニュース記事の要約作業を自動化するシステムのプロトタイプを開発しました。今後実用化をめざし、日本テレビ内にて現場担当者による事前の性能テストを行います。

ドコモは2016年より、AIを活用した本システムの研究開発に着手しており、文章の中から重要度の高い文章をAIが判定して抜粋したり、AIが言葉を補いながら新たに文を生成して要約したりする本システムを作成していました。

両者は2018年9月から、日本テレビが大量に保有する20万件近い過去記事データと人が要約したデータセットを本システムに学習させ、要約の精度向上に取り組むとともに、2019年5月から開始した本実証実験では、本システムの機能拡充やインターフェースの実装を行い、現場で使用するためのWEBアプリケーション型のプロトタイプを開発しました。

今回の機能拡充においては、日本テレビの原稿制作ノウハウをルールベースで実装するとともに、タイトルに含まれる単語やユーザーが指定した任意の単語の重要度を向上させ要約結果に反映する機能や、要約後の単語が要約前文章のどの部分で出現しているかをトレースできる機能など、現場の声を反映したアイデアを追加実装しています。

本システムの基盤となる技術には、NTTドコモ北京研究所(以下、北京研)の自動要約技術を用いています。本要約技術は、北京研がAI技術を用いて独自に開発したニューラルネットワーク(DNN:Deep Neural Networks)^{*1}により実現されており、既存技術と比較して重要情報の網羅性を高めただけでなく、必要な接続詞などを選択・生成することにより読みやすい文章を生成します。

なお、本実証実験で開発したプロトタイプについて、2020年1月23日(木)～24日(金)に開催するドコモのイベント「DOCOMO Open House 2020」のAI展示ブースにて、実際にデモをご体感いただけます。

両社は今後現場での検証を通じてさらなる要約精度の向上を図り、本システムの実用化をめざしてまいります。

※1 ニューラルネットワーク:人間の脳の仕組みを模したAIの計算モデル手法のこと。

本件に関するお問い合わせ先	
日本テレビ放送網株式会社 社長室広報部 TEL: 03-6215-3090	株式会社NTTドコモ 第一法人営業部 第二営業担当 TEL: 03-5156-2340

実証実験概要

■ 背景

日本テレビが運営するニュース専門サイト「日テレNEWS24」では、最新のニュースを動画とニュース原稿(全文記事と要約記事)で配信しており、特に要約記事をサイト上に表示することにより読者にニュース概要を分かりやすく伝えています。一方で、このニュース記事の要約作業には人手による要約の手間が発生しています。また、要約作業には熟練したスキルが必要であるため、スタッフの育成に一定期間の研修が必要で、人員の確保も課題となっています。

■ 本システムの概要

本システムは、要約手法の異なる「抽出式要約システム」と「生成式要約システム」の2つの自動要約システムに大別され、各々についてプロトタイプを開発しました。

1. 抽出式要約システム

抽出式では、全文原稿から重要度の高い文を抜き出して要約します。

<基本機能>

- ・ AIを用いた文書要約機能
- ・ 要約元の全文原稿から、文単位でそのまま文章を抽出し、要約結果として表示

◆通常の抽出式自動要約

全文原稿

台風の影響を受け東京は大雨。

過去最大の降水量を記録した。

交通機関への影響が出ている。

今後も注意が必要だ。



要約

台風の影響を受け東京は大雨。

交通機関への影響が出ている。

<追加機能>

- ・ 要約元原稿のニュース記事のタイトルやユーザーが指定した任意の単語を重要語として判定し、要約結果に反映する機能
- ・ 要約結果が要約元原稿のどの文章から抽出されたのかをトレースする機能
- ・ あらかじめ設定した上限文字数内で要約を生成する機能

◆ヒントありの抽出式自動要約

全文原稿

台風の影響を受け東京は大雨。

過去最大の降水量を記録した。

交通機関への影響が出ている。

今後も注意が必要だ。

人間が重要部分
を選択

要約原稿（ヒントあり）

台風の影響を受け東京は大雨。

過去最大の降水量を記録した。

→人間が全文テキストの中から任意の重要部分を選択した場合、その重要部分が含まれる文を優先的に抽出することで、重要部分が必ず要約テキストに含まれるようにします。

2. 生成式要約システム

生成式では、全文原稿からの抽出だけでなく、言葉を補いながら新たに文を生成して要約します。

<基本機能>

- ・ AIを用いた文書要約機能
- ・ 要約元の全文原稿を参考に、重要な文を抜き出して新たな文書を生成

◆通常の生成式自動要約

全文原稿

台風の影響を受け東京は大雨。

過去最大の降水量を記録した。

交通機関への影響が出ている。

今後も注意が必要だ。

要約

台風で東京は大雨。

交通機関への影響に注意が必要だ。

<追加機能>

- ・ 要約元原稿のニュース記事のタイトルやユーザーが指定した任意の単語を重要語として判定し、要約結果に反映する機能
- ・ 要約結果が要約元原稿のどの文章から生成されたのかをトレースする機能
- ・ あらかじめ設定した上限文字数内で要約を生成する機能
- ・ 起承転結を考え、必要な接続詞などを選択・生成する機能

◆ヒントありの生成式自動要約

全文原稿

台風の影響を受け東京は大雨。
過去最大の降水量を記録した。
交通機関への影響が出ている。
今後も注意が必要だ。



要約原稿（ヒントあり）

台風で東京は過去最大の降水量を記録。
交通機関への影響に注意が必要だ。

人間が重要部分
を選択

→人間が全文テキストの中から任意の重要部分を選択した場合、その重要部分に含まれる単語が必ず要約テキストに含まれるようにするとともに、その前後の文脈的に必要な部分の重要度を高め、結果的に要約文に含まれやすいようにします。