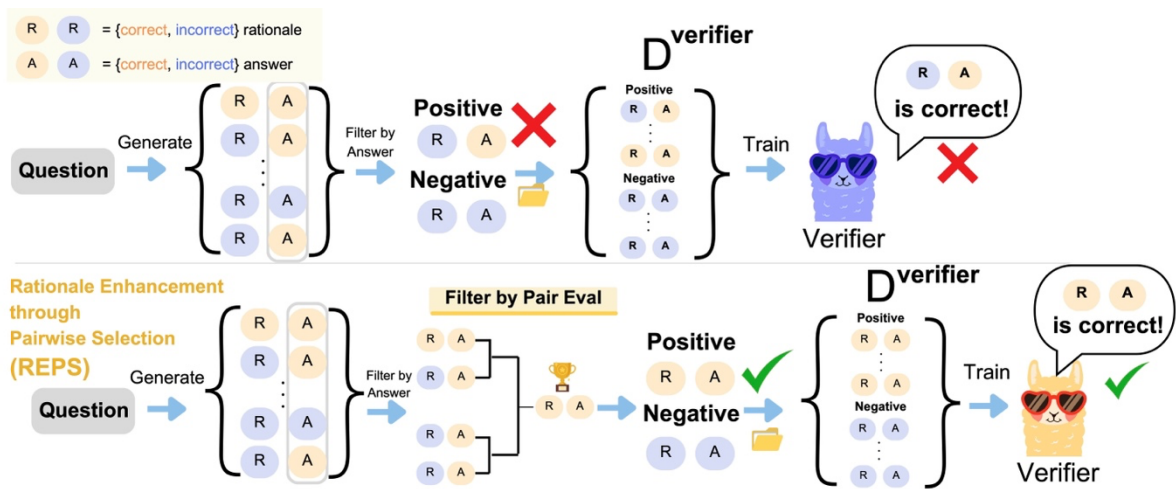


〈報道関係のみなさま〉

国際学会「EMNLP」で本社員の論文が 2 年連続で採択 大規模言語モデルの信頼性向上に関する手法を提案

株式会社朝日新聞社(代表取締役社長 角田 克)メディア研究開発センター員の川畑輝が主著した論文(以下「本論文」)が、自然言語処理分野における世界最高峰の国際学会の一つ「Empirical Methods in Natural Language Processing」(以下「EMNLP」)の本会議に採択されました。川畑が主著の論文が EMNLP に採択されるのは 2023 年に続き 2 年連続になります。本年は 11 月 12 日から 16 日まで米国フロリダ州マイアミで本会議が開催され、現地で川畑が発表します。

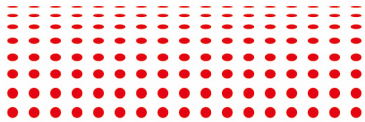
本論文は、国立情報学研究所 菅原朔助教との研究成果をまとめたもので、大規模言語モデル(以下「LLM」)が応答する内容の、真偽性を自動評価するための LLM(以下「評価用 LLM」)の構築手法について提案しました。従来研究において評価用 LLM は、評価対象の LLM が出力する最終的な結論の真偽のみに評価の焦点を当て構築されてきました。本論文では、評価対象の LLM が出力する結論だけでなく、結論を導くまでの過程や根拠の妥当性も評価できるような評価用 LLM を学習する手法「Rationale Enhancement through Pairwise Selection:REPS」を提案しました。情報の受け手にとっては「結論のみならず結論を導くまでの過程や根拠も重要である」ことに着目したもので、LLM 自体の性能向上や、LLM の応答に含まれる嘘(ハルシネーション)の低減や検知にもつながる研究です。



REPS の概念図(本論文より引用)



〈お問い合わせ〉
朝日新聞社 メディア事業本部
メディア研究開発センター
E-mail mrاد-contact@asahi.com



press release

朝日新聞

朝日新聞社 〒104-8011 東京都中央区築地 5-3-2

朝日新聞社は最新テクノロジー、特にAI分野での研究を推進し、社内外のLLM技術の利活用やDXに貢献していきます。

■論文について

Akira Kawabata, Saku Sugawara. Rationale-Aware Answer Verification by Pairwise Self-Evaluation. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP 2024)*, Florida, U.S.A., November 2024. <https://arxiv.org/abs/2410.04838>

邦題: ペアワイズ自己評価による根拠を考慮した回答検証

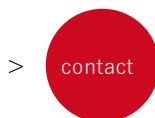
■メディア研究開発センターについて

2021年4月に発足しました。人工知能を始めとする先端メディア技術と、新聞社ならではの豊富なテキストや写真、音声などの資源を活用し、社内外の問題解決を目指すとともに、自然言語処理や画像処理をはじめとした先端技術の研究・開発を進めています。

<本リリースについての問い合わせ>

朝日新聞社 メディア事業本部 メディア研究開発センター

E-mail: mrad-contact@asahi.com お願いします。



<お問い合わせ>

朝日新聞社 メディア事業本部
メディア研究開発センター

E-mail mrad-contact@asahi.com