

3. 記述&対話 AI

2023年度までの進捗状況

1. 概要

研究者も普段は実験結果をまとめて他の研究者と議論し、さらなる仮説を立てて次の実験に入っていきます。このように結果を記述し、対話的に議論することで仮説をアップデートできるような AI が必要となります。

このプロジェクトでは、(1)仮説として実験をデザインする「Scientist-in-the-loop による論文理解・実験計画 AI」と(2)対話的に仮説を立てていく「知識推論と対話を用いたマルチモーダル仮説生成」の研究開発を進め、上記のような記述&対話 AI を実現します。

2. これまでの主な成果

論文中の図表と文から実験内容を理解するマルチモーダル XAI の開発を目指し、研究者の経験を反映させる技術を導入しました。研究者のフィードバックを活用し、大規模データセットなしで LLM の性能を向上させ、専門性に応じたタスク割り当てとプロンプト信頼性を高める技術を開発しました。

さらに、マルチモーダル仮説生成の基盤として大規模言語モデルを用いた論文理解と仮説生成モデルの構築に取り組みました。

(1)論文理解・実験計画 AI

まず、マルチモーダル XAI の実現に向けて、LLM に研究者の暗黙知を反映させる技術を開発しました。研究者の具体的な出力例を使い、プロンプトを通じて知識を取り入れる In-context learning を活用し、関連記述の抽出とその根拠の出力を行う技術を開発しました。これにより、大規模データセットを必要とせずに LLM の性能向上を実現しました。

次に、研究者の知識理解 AI を開発し、専門性に応じたフィードバックタスクの最適な割り当てとプロンプトの信頼性を推定



図1 研究者からの知識理解 AI

する技術を開発しました。これにより、複数の研究者が協力して LLM のプロンプトを設計する際に、質の高い出力を得ることができました(図1)。

さらに、研究者からの知識獲得 AI と知識理解 AI による実験内容理解に取り組みました。物体検知技術に関する論文を対象に、タグの抽出と検証を行い、LLM を単純に利用する場合と比べて精度が向上しました。研究者の知見を取り入れることで、実験内容の理解が効率化されることを確認しました(図2)。



図2 実験内容理解に基づいたタグ抽出と関連文献の提示

最後に、分子合成 AI への研究者の知見を埋め込むためのインターフェースとして、分子編集のウェブアプリを開発しました。RLHF 技術を利用し、研究者が基盤モデルの出力を評価するだけでなく、改善案を提案することで、研究者の思考をモデルに反映させることができました。

(2)マルチモーダル仮説生成

この研究では、知識推論と対話を用いたマルチモーダル仮説生成の課題に取り組みました。まず、大規模言語モデルを用いた論文理解枠組みの構築に注力し、特許庁から取得した論文データベースを用いて、既存の open-calm と T5 モデルをベースに調整しました。目的関数には一般的なマスクランゲージモデルを使用しました。ファインチューニングされたモデルにプロンプトを与え、仮説生成を行いました。特許データから因果関係に関わる 10 万件の仮説生成データを作成しました。

仮説生成に必要なデータ整備も行い、言語学に精通したアナデータを雇用して、特許データ 10,000 件に対してアノテーション標準を構築し、選択した特許に 300 以上のアノテーションを行いました。

知識推論モデルの構築では、因果関係に焦点を当て、生成した仮説データと外部の因果関係辞書を用いて妥当な仮説を選別する方法を検討しました。仮説について研究者と対話する AI の構築に向けて、強化学習を用いた AI 発話修正方法(RL-AIF)の検討を行いました。今後、この手法の評価を行う予定です。

3. 今後の展開

引き続き研究者の対話から研究者のフィードバックを通じて AI の性能を向上させる共進化的な人工知能技術を開発し、仮説生成にむけた研究者のインスピレーションの AI への埋め込みを目指します。仮説インスピレーション AI が知識を獲得・推論し、人間と協調して対話を行うことで、革新的な発見とイノベーションを引き起こします。