



XLLMN: 外部検索キーワードネットワーク可視化によるLLM出力のリアルタイム説明手法

杉本拓海 (先進情報/1年) 北條湊都 (社会情報/1年)
秋山知輝 (先進情報/1年) 木村翔竜 (社会情報/1年)
山田祐平 (人工知能/4年) ソン・チェヨン (人工知能/4年)
西井正平 (先進情報/1年) 高山隆司 (先進情報/1年)

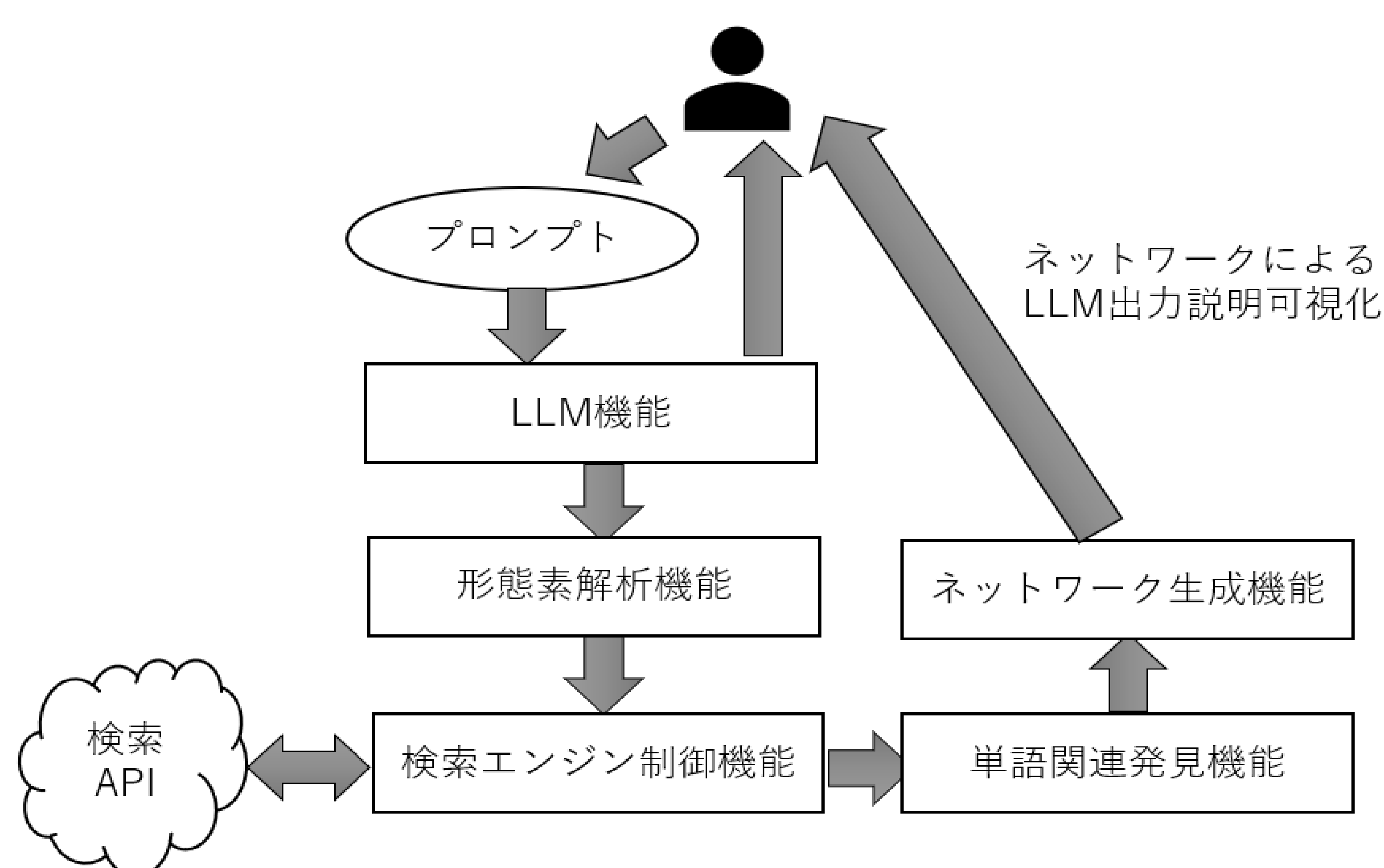
概要

現在、生成AIやLarge Language Model(LLM)の進展により、多くのユーザがAI、機械学習を気軽に使う機会が増大した一方で、AI、機械学習を用いた意思決定においては、AI、機械学習モデルは根拠や説明を導出することが難しいブラックボックスモデルであることが多く、出力の根拠を示すことが難しいため、人命や財産に関わる意思決定にAI、機械学習を使う機会が減少している事実がある。特にLLMについては、相関関係で出力されることが多く、Hallucinationの問題がある。本研究では、外部検索キーワードネットワーク可視化によるLLM出力リアルタイム説明手法について示す。本研究は、ユーザが入力したプロンプトに対し、LLMから出力された文章をキーワードに分割し、そのキーワードで外部検索エンジンを使い上位の検索結果のコンテンツをスクレイピングし、ユーザが入力で使った言葉を探査し、LLMの出力の単語とユーザの入力の単語のネットワークを構成することで、LLMの説明可能性を高める。本方式によって、なぜLLMがユーザのプロンプトに対して出力したかの説明を可視化することが可能となる。

背景・目的

- ・ 昨今、現代社会において、AIが活用されている事例が多く存在するが、その事例の多数は、入力と出力のみで過程についてはブラックボックス化している。
- ・ そこでAIが結果を出力する際に、その導出過程を可視化することができる。AI自身が出力の説明を行い、AIによる思考過程の可視化を目的とした研究である。

全体図



本方式のポイント

- ・ Gemini APIを利用した思考プロセスの生成と最終回答の提案。
- ・ 思考プロセスは、問題定義、仮説提示、情報収集、分析、結論などの
- ・ ステップで構成される。
- ・ Google Custom Search APIを利用したWeb検索機能。
- ・ LLMからの応答の解析と構造化。
- ・ 思考プロセスをネットワーク図として可視化。

実験(出力例)



まとめ・今後の課題

- ・ 本研究では、外部検索キーワードネットワーク可視化によるLLM出力リアルタイム説明手法について示した
- ・ 本方式について、ネットワーク形式でLLMの導出理由を可視化するシステムを構築した
- ・ Fake Newsの検知や。審議の判断への適用
- ・ LLM自体の性質(グロッキング)の科学的な検証
- ・ LLMの有効な活用のためのプラットフォーム形成