

情報の真偽を検証するプログラム

ライセンス契約を受けていただき 本発明の実用化を目指していただける企業様を求めます。

テキストデータに含まれる主張内容の信憑性を高精度で算出するプログラムです

◆背景

オンラインソーシャルメディア（Web掲示板、ブログ、SNS等）は、多くのユーザが情報を求めたり、情報を広めたり、審議や議論を行ったりする新しい手段となっています。このため、オンラインに書き込まれた噂や誤報が、以前よりも大きな影響力を持つようになりました。したがって、オンラインの情報の信憑性を検証することが重要な課題となっています。

◆発明概要と利点

本発明は、テキストデータに含まれる主張内容の信憑性を高精度で算出するプログラムです。具体的な判定フローは、既存の大規模言語モデル（LLM）を用いてテキスト情報から主張を抽出し、当該主張の真偽を確認する質問（Query）を生成して複数のエージェント（例えばWikipedia, Google等の検索サイト）で検索結果を取得します。その後、当該検索結果と当該主張とを比較して得られる各エージェントでの真偽性の判定結果（真／偽、および確信度）を集約し、独自に、真の確信度と偽の確信度を踏まえてその加重平均を総合判定に用いて情報の信憑性を算出します。

- 多様な観点からのファクトチェック（真偽チェック）を行うことができる。
- 個々のエージェント（検索サイト）を用いて検証を行った場合と比べ精度が高い（表1）。
- 複数のエージェント（検索サイト）の結果を用いた多数決よりも精度が高い（表1）。
- シミュレーション、SNS情報、過去の事例情報、ナレッジグラフなどの様々なエージェント（検索サイト）を用いることができる。

◆研究段階

- ・単一エージェント（検索サイト）、複数のエージェント（検索サイト）の結果を用いた多数決、および本発明の性能を比較済み。
- ・GPT-3.5-turboと本発明の性能を比較済み。

◆適応分野

- ファクトチェック（真偽チェック）**
- ・オンラインソーシャルメディアの情報管理
 - ・会議、意見集約時の情報管理

◆希望の連携形態

- ・実施許諾契約
 - ・プログラムのライセンス
 - ・オプション契約（技術検討のためのF/S）
- ※1：本発明は京都大学から特許出願中です。
- ※2：情報処理学会で発表済み

◆お問い合わせ先

京都大学産学連携担当
株式会社TLO京都

〒606-8501
京都市左京区吉田本町
京都大学 産官学連携本部内
(075)753-9150
event@tlo-kyoto.co.jp

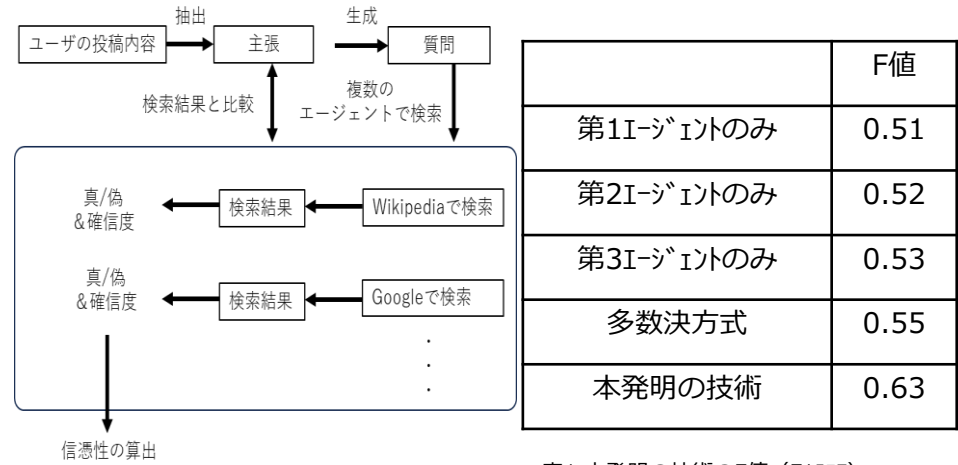


図1.本発明の判定方式フローチャート

表1.本発明の技術のF値（F1スコア）
F値が高いほど信頼度が高い