

記号推論に接続する機械学習

研究代表者 杉山 鷹人 (国立情報学研究所 情報学プリンシプル研究系 准教授)



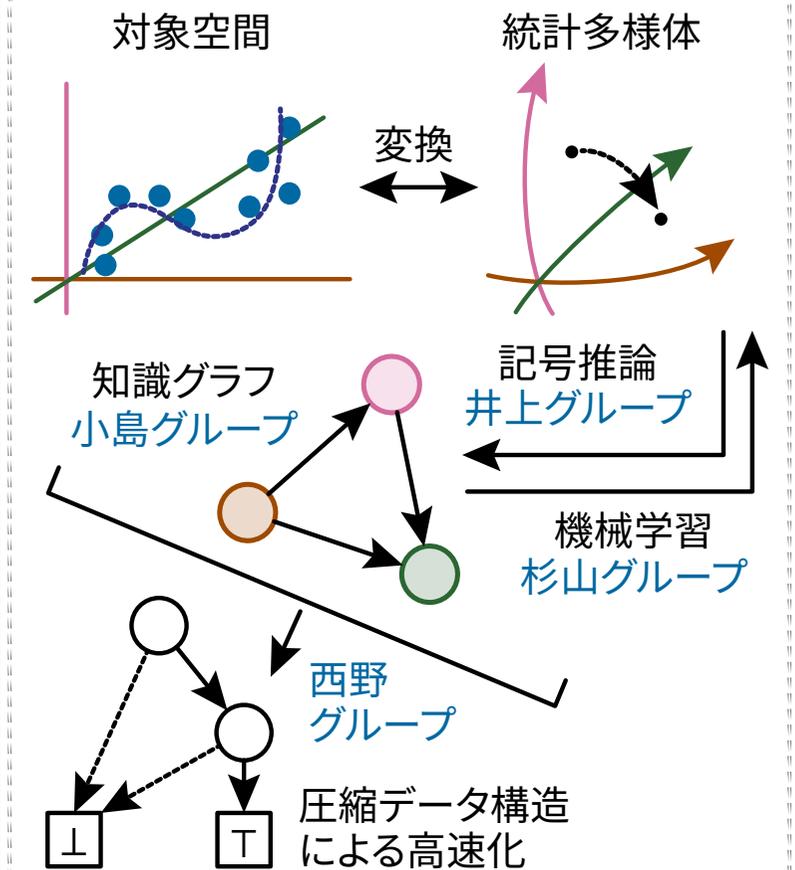
研究領域「信頼されるAIシステムを支える基盤技術」
(研究総括: 相澤 彰子、2020年度発足)

研究の概要: 信頼されるAIシステムの基盤となる基礎技術の創出を目的として、主に幾何学的なアプローチに着目することで、現代的な機械学習と推論根拠の解釈性に優れた記号推論の融合を実現する。記号推論を前提とした機械学習システムを設計・構築することで、機械学習がもつ信頼性についての課題と、記号推論が持つロバスト性についての課題を同時に解決する。特に、大量パラメータを用いた学習による高い汎化能力と、連続ドメインに直接適用可能な推論による説明性・解釈性の担保を両立する。

研究終了時の達成目標: 高い汎化能力と連続ドメインに適用可能な推論を両立し、学習と推論がスムーズに連結された基盤を確立する。

研究の独創性、新規性・優位性

説明性に着目するXAIやソフトウェア工学の技術を応用する機械学習工学、深層学習と記号推論の統合を目指すニューロ・シンボリックAIの研究が国内外で進んでいる。これらの既存研究では、いかに機械学習の内部に触れずブラックボックスのまま信頼性を担保するのか、という方針をとる。これに対して本提案研究は、記号推論が直接実行可能な機械学習方式を再設計することで、ホワイトボックス性を保ったまま、大量パラメータを利用した高い汎化性能とスケーラビリティの達成を目指す独創的な試みである。このアプローチは新規であり、モデルの論理的解釈が本質的に可能という優位性を持つ。



研究の将来展望:

(1) CREST研究期間終了後の研究計画

機械学習の技術そのものの発展や実社会での応用は、本提案研究の実施中にも急速に進むことが予想される。そこで、本提案研究で構築したAIシステムを出発点として、それら最先端の技術を取り込むための基礎技術から、サイエンス領域・実社会への応用まで、幅広く研究を展開する。

(2) 科学技術イノベーション(※)創出、知的財産権の取得・活用、新産業創出・社会貢献

本提案研究により、データサイエンスのさらなる利活用が促進されることが期待される。したがって、データサイエンスに基づく新産業や社会貢献を促進することが期待される。また、方法論研究が主体となるため、知的財産権の取得は予定せず、システムをオープンソースで公開する。

(※)「科学的な発見や発明等による新たな知識を基にした知的・文化的価値の創造と、それらの知識を発展させて経済的、社会的・公共的価値の創造に結びつける革新」出典: 第4期科学技術基本計画