

信頼される AI システムを支える基盤技術  
2022 年度採択研究代表者

2022 年度  
年次報告書

杉山 磨人

情報・システム研究機構 国立情報学研究所  
准教授

記号推論に接続する機械学習

主たる共同研究者:

井上 克巳 (情報・システム研究機構 国立情報学研究所 教授)

小島 諒介 (京都大学 大学院医学研究科 講師)

西野 正彬 (日本電信電話(株) NTT コミュニケーション科学基礎研究所 特別研究員)

## 研究成果の概要

各グループが本プロジェクトの目標である機械学習と記号推論の接続を念頭に置きつつ研究に取り組み、それぞれの領域で成果を挙げることに成功した。まず杉山がリーダーを務める機械学習グループでは、記号推論と相性の良い機械学習手法である決定木の理論解析に取り組み、NTK (Neural Tangent Kernel) と呼ばれる枠組みを用いた決定木アンサンブルの解析に成功し、非同型だが学習挙動が同じになる木構造の存在などいくつかの非自明な性質を明らかにした[1]。さらに、西野がリーダーを務めるアルゴリズムグループでは、学習モデルに検証器を接続した際の挙動を理論的に解析し、学習時にも検証器を用いることで汎化性能の上界が検証内容に依存しないことを Rademacher 複雑性を用いて示した[2]。また、機械学習グループにおいて、代表的なデータ形式である行列・テンソル表現に対して、情報幾何学を用いることで特徴間の関連を直接扱う統計的モデリング技術を構築した[3]。井上がリーダーを務める記号推論グループでは、機械学習で得られる連続系のデータと記号推論を自然な形で接続しベクトル空間での論理推論を実現するために、記号推論を行列・テンソル表現し線形代数演算と連続最適化を用いて実現するための理論的基礎及び実践的手法の開発を進めている。また機械学習モデルの入出力関係を正しく説明し、状況変化に適応したロバストなモデルとするために、知識表現・推論における信念修正の研究で提案されていた合理的基準を分類器の編集に適用し、編集オペレータに対する構成的特徴付けを与えた[4]。同時に、小島がリーダーを務めるモデリンググループでは、記号推論をテンソル演算で実現する確率論理プログラミング言語 T-PRISM の開発を進め、さらに時系列データから入出力安定性を保証しつつ非線形システムを学習するロバストな手法を構築した[5]。

### 【代表的な原著論文情報】

- 1) Ryuichi Kanoh, Mahito Sugiyama, “Analyzing Tree Architectures in Ensembles via Neural Tangent Kernel”, 11th International Conference on Learning Representations (ICLR 2023), 2023.
- 2) Masaaki Nishino, Kengo Nakamura, Norihito Yasuda, “Generalization Analysis on Learning with a Concurrent Verifier”, Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS 2022), 2022.
- 3) Kazu Ghalamkari, Mahito Sugiyama, “Non-negative Low-rank Approximations for Multi-dimensional Arrays on Statistical Manifold”, Information Geometry, 2023.
- 4) Nicolas Schwind, Katsumi Inoue, Pierre Marquis, “Editing Boolean Classifiers: A Belief Change Perspective”, 37th AAAI Conference on Artificial Intelligence (AAAI 2023), 2023.
- 5) Ryosuke Kojima, Yuji Okamoto, Learning Deep Input-Output Stable Dynamics, Thirty-sixth Conference on Neural Information Processing Systems (NeurIPS 2022), 2022.