

研究課題別事後評価結果

1. 研究課題名： ロバスト性と汎化性能を両立する機械学習法の確立

2. 個人研究者名

藤澤 将広（理化学研究所革新知能統合研究センター 基礎科学特別研究員）

3. 事後評価結果

本研究では、データに敵対的摂動や外れ値が含まれる場合を想定して、汎化性能とロバスト性を同時に実現する機械学習法の確立を目標とした。しかし、研究を進める過程で、汎化性能に関する現状の理論ではロバスト性との議論が困難であることが判明したため、この困難を克服し得る汎化理論の提供を目指すことに研究の方向性を修正した。最初に、確率的勾配 Langevin 動力学に対して、訓練データの安定性の時間発展を追跡することで、新たな情報理論的汎化誤差上界を導出した。また、Stein 変分勾度降下法(SVGD)を導入し、無限粒子および有限ステップ幅の下での SVGD が、KL 距離基準で準線形収束することを示した。これらの研究成果のうち、前者は機械学習分野のトップカンファレンスである NeurIPS2023 に於いて発表され、高く評価されている。後者についても、主要な国際会議に投稿された。また、国際会議 BayesComp2023 では招待講演を行っており、世界的に多くの関心を集めている。今後の課題として、汎化性能と事前分布の誤設計に対するロバスト性を両立する新たな学習アルゴリズムを検討しており、更なる発展が期待される。