

信頼される AI の基盤技術
2022 年度採択研究代表者

2022 年度
年次報告書

二見 太

大阪大学 大学院基礎工学研究科
講師

情報理論を用いた不確実性に関する学習理論の展開

研究成果の概要

本年度は研究計画 A)「情報理論を使った不確実性の定量化および学習理論の構築」に該当する箇所の研究を行った。具体的には、まず変分ベイズ推論において、典型的に使われる目的関数に対し、数学的に目的関数を予測性能とばらつきに関する新しい最適化問題として再立式し、それを PAC ベイズの手法で解析を行った。この解析では、学習データの不足に由来する損失を「モデルがまだ何を学習できていないか」を表現する不確実性としてとらえることを提案した。これによりこれまで既存の変分ベイズでは学習の目的関数の選択により異なる不確実性を出力する解が選択されることが知られていたが、本研究では、そうした目的関数の選択が学習データの不足に由来する損失に与える影響を解析することで、異なる不確実性がもたらされる理由を解明した。

次に、研究計画に記した、情報理論を使った解析手法として、学習データと学習したモデルに関する相互情報量に着目する新たな解析法を考案し、Stochastic gradient Langevin dynamics に適用することを行った。この方法は PAC ベイズの方法で解析が困難な「具体的なアルゴリズム」や事前分布の選択の問題点について情報理論の道具を活用することで解決を試みるアプローチである。

最後に、具体的なアルゴリズムではなく、特定の条件が満たされた場合にアルゴリズムの予測の不確実性がどのようにふるまうのか解析を行った。具体的には、学習において訓練データに対する最適解が得られた場合かつモデルの選択ミスが発生しない場合など、学習において「ある種の理想的な条件」が満たされた場合でかつ学習データとテストデータが独立同分布に従う場合に、予測の不確実性がどのようにふるまうのかを解析したものである。これによって不確実性としてどのような振る舞いが統計的学習理論の枠組みで望ましいのかその方向性を明らかにした。

これらの3つの解析結果について現時点で別々に論文執筆中であり、2023 年度中に国際会議及び論文誌に投稿予定である。