

情報理論を用いた不確実性に関する学習理論の展開

基礎工学研究科 システム創成専攻

准教授 二見 太

Researchmap <https://researchmap.jp/ffutami>



研究の概要

機械学習の応用先が広がるにつれ、単に高い予測性能を達成するためだけに確率的推論を用いるのではなく、予測の不確実性の定量化や、潜在変数モデルによる知識の抽出・生成など、多様な活用が進んでいる。しかし、これらの手法に対する理論的理解は十分ではない。本研究では、このような機械学習アルゴリズムを情報理論に基づく手法で解析し、予測の不確実性や潜在変数の振る舞いに対する統計的学習理論を体系的に構築することを目指す。

在変数モデルが、情報理論的な汎化誤差解析の枠組みに自然に組み込めることを明らかにした。従来の学習理論では、潜在変数そのものを解析対象とすることが難しく、主にエンコーダのパラメータを介してモデルの複雑さを評価していた。本研究で確立した枠組みは、潜在変数を直接的な理論対象として扱うことを可能にし、深層生成モデルに対する学習理論の適用範囲を大きく拡張した点で重要な成果である。

研究の意義と将来展望

本研究は、機械学習の信頼性を支える上で不可欠な要素であるアルゴリズムの予測やモデルに内在する不確実性に着目し、その適切な定量化と理論的体系の構築を目的とする。不確実性は、予測のばらつき、信頼区間、クラス事後確率、さらにエンコーダ型モデルによって得られる潜在変数表現などの形で現れ、意思決定において中心的な役割を果たし、AIの安全性および信頼性の確保に直結する。しかし既存の理論は、分類精度や二乗誤差といった損失最小化に主眼が置かれており、不確実性に対する理論的保証や潜在変数の挙動に関する理解は十分に確立されていない。本研究では、情報理論とPACベイズ理論を融合することで、不確実性評価に焦点を当てた新しい統計的学習理論を構築し、将来的には高信頼なAIシステムの設計指針へと発展させることを目指す。

研究の背景と結果

確率的推論に基づくアルゴリズムは、ベイズ推論に支えられた一貫した学習原理の下で発展してきた。現在では、深層学習と組み合わせることで、潜在変数を用いた知識の抽出・生成、不確実性評価、またそれらを用いた意思決定などに対して強力な学習規則を提供している。一方で、現代の確率的推論は必ずしもベイズの定理そのものに限定されず、変分ベイズを中心とした近似推論や、一般化ベイズによる柔軟な更新則など、問題設定に応じた多様なアルゴリズムが提案されている。

しかし、このような「ベイズの更新則に必ずしも限られない」確率的推論に対する学習理論は、これまで主に独立同分布データを仮定し、期待誤差を対象とする汎化誤差解析にとどまってきた。潜在変数モデルによる生成・表現学習、不確実性の精密な定量化といった現代的な応用に対しては、理論的理解が十分でなく、その結果として推論アルゴリズムの設計やモデル評価の多くがヒューリスティクスに依存している。そのことは、高不確実性環境におけるモデルの汎用性の制限や、予測結果の信頼性の低下といった問題につながっている。

本研究では、従来は経験的に扱われてきた不確実性および深層学習を用いたエンコーダによる隠れ変数の評価を理論的に定式化及び保証することに成功した。まず、情報理論と古典的統計的学習理論の対応関係を整理し、情報量や損失関数の滑らかさなどを通じて、モデルの一般化性能と不確実性を統一的に議論できることを示した。

さらに理論解析の過程で、変分オートエンコーダのような深層潜

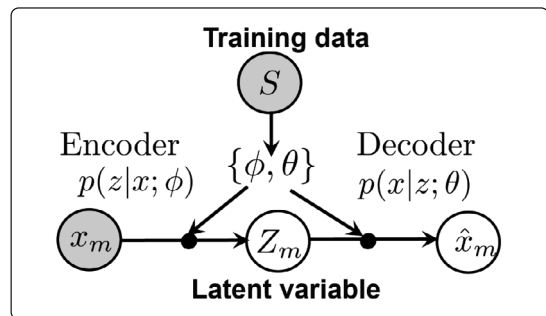


図1: Encoderを用いた隠れ変数モデル(変分オートエンコーダなど)

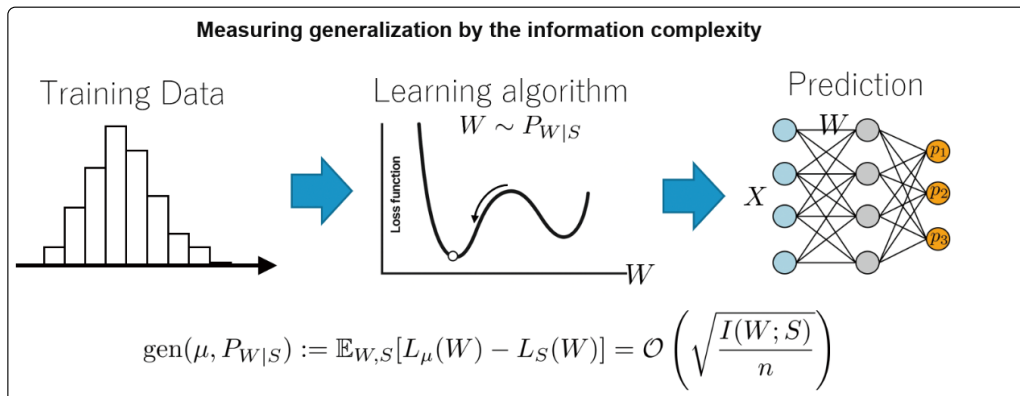


図2: 機械学習における統計的学習理論と情報理論の関係図

特許

Futami, Futoshi & Fujisawa, Masahiro. Information-theoretic Generalization Analysis for VQ-VAEs: A Role of Latent Variables. Accepted at Advances in Neural Information Processing, 2025.
 Fujisawa, Masahiro & Futami, Futoshi. PAC-Bayes Analysis for Recalibration in Classification. Proceedings of 42nd International Conference on Machine Learning, 267, 17986-18023, 2025. doi: 10.48550/arXiv.2406.06227
 Futami, Futoshi. Epistemic Uncertainty and Excess Risk in Variational Inference. Proceedings of 28th International Conference on Artificial Intelligence and Statistics, 258, 568-576, 2025.
 Futami, Futoshi & Fujisawa, Masahiro. Information-theoretic Generalization Analysis for Expected Calibration Error. Advances in Neural Information Processing Systems 37, 84246-84297, 2024. doi: 10.48550/arXiv.2405.15709
 Futami, Futoshi & Iwata, Tomoharu. Information-theoretic Analysis of Bayesian Test Data Sensitivity. Proceedings of 27th International Conference on Artificial Intelligence and Statistics, 238, 1099-1107, 2024.

参考URL <https://sites.google.com/view/futoshi-futami/home>

キーワード 統計的学習理論、機械学習、生成AI、情報理論、ベイズ推論