

理化学研究所
革新知能統合研究(AIP)センター
汎用基盤技術研究
グループの取り組み

グループディレクター 杉山 将

汎用基盤技術研究グループの目標

- 学習理論・最適化の**原理解明**
- 新しい学習・推論**アルゴリズムの開発**
- 機械学習における国際的な「**ハブ**」機能の確立

汎用グループの体制

13チーム・ユニット
常勤研究員 32 人
(2023年12月時点)

実応用

新しい学習・推論アルゴリズムの開発

テンソル学習チーム

無期PI



チームリーダー
Qibin Zhao (Ph.D.)

不完全情報学習チーム



チームリーダー
杉山 将 (Ph.D.)

近似ベイズ推論チーム

無期PI



チームリーダー
Mohammad Emtiyaz Khan (Ph.D.)

逐次的意思決定チーム

新設



チームリーダー
伊藤 伸志 (Ph.D.)

因果推論チーム



チームリーダー
清水 昌平 (Ph.D.)

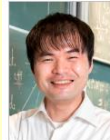
構造的学習チーム



チームリーダー
河原 吉伸 (Ph.D.)

圧縮情報処理ユニット

常勤



ユニットリーダー
田部井 靖生 (Ph.D.)

計算論的学習理論チーム



チームリーダー
畑 晃平 (Ph.D.)

高次元因果解析チーム

新設



チームリーダー
今泉 允聡 (Ph.D.)

理論保証

学習理論・最適化の原理解明

深層学習理論チーム



チームリーダー
鈴木 大慈 (Ph.D.)

連続最適化チーム



チームリーダー
武田 朗子 (Ph.D.)

関数解析的学習チーム

常勤



チームリーダー
Quang Minh Ha (Ph.D.)

数理科学チーム



チームリーダー
坂内 健一 (D.Math.Sci.)

汎用グループの組織改編と強化

- 効率的なマネジメントおよび連携の強化のため、規模の小さい10チームを**統合**（2020年度、2021年度）

21チーム/ユニット→11チーム/ユニット

統合したチーム

- オンライン意思決定ユニット
- 離散最適化ユニット
- 幾何学的学習チーム
- 高次元統計モデリングチーム
- 数理解析チーム
- 非凸学習理論チーム
- マルチエージェント最適化チーム
- 数理統計学チーム
- トポロジカルデータ解析チーム
- 探索と並列計算ユニット

- AIP の強みである機械学習の技術開発とそれを裏付ける原理解明研究を**強化**（2023年度）

高次元因果解析チーム、逐次的意思決定チームを新設



代表的な研究成果(深層学習理論)



鈴木大磁ら

深層学習が浅層学習よりも優れた予測性能を持つことを証明

成果発表 NeurIPS (2020, 2021), ICLR (2021)

概要

機械学習では高次元かつ複雑なデータを用いた学習が要求されるが、深層学習におけるパラメータの最適化は非常に困難。

本成果では、深層学習では大域的最適化と高次元予測が可能であることを数学的に証明。

(大域的最適化が可能)

$$\hat{L}(X_k) - \int \hat{L}(x) d\pi_\infty(x) \lesssim \exp(-\Lambda_\eta^* k \eta) + \frac{C_\beta}{\Lambda_0^*} \eta^{1/2-\kappa}$$

$$dX_t = -\nabla \left(\hat{L}(X_t) + \frac{\lambda}{2} \|X_t\|_{\mathcal{H}_K}^2 \right) dt + \sqrt{\frac{2}{\beta}} d\xi_t$$

(高次元予測が可能)

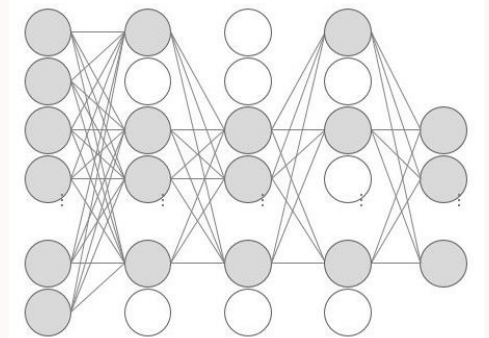
$$R_{\text{lin}}(\mathcal{F}_\gamma) \gtrsim n^{-\frac{2\tilde{\beta}+d}{2\tilde{\beta}+2d} - \kappa'}$$

$$R_{\text{lin}}(\mathcal{F}_\gamma) := \inf_{\hat{f}: \text{linear}} \sup_{f^\circ \in \mathcal{F}_\gamma} \mathbb{E}_{D_n} [\|\hat{f} - f^\circ\|_{L_2(P_X)}^2]$$

$$\mathbb{E}_{D^n} \left[\mathbb{E}_{W_k} [\|f_{W_k} - f^\circ\|_{L_2(P_X)}^2 | D_n] \right] \lesssim n^{-\frac{\gamma}{\alpha_1 - 3\alpha_2 + 1}} + \Xi_k$$

成果の意義・社会貢献への期待

本成果はICLR2021にてOutstanding Paper Awardを受賞。深層学習の更なる性能向上のための基盤となり、モデルの圧縮、説明性やロバスト性の向上など、信頼できるAIの開発に貢献できると期待。



代表的な研究成果(弱教師付き学習理論)



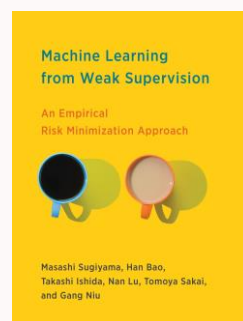
杉山 将ら

様々な弱教師付き学習問題に適用できる汎用的な学習理論とアルゴリズムを構築

成果発表 ICLR2022, AISTATS2022, ICML2021, NeurIPS2020, ICML2020 他

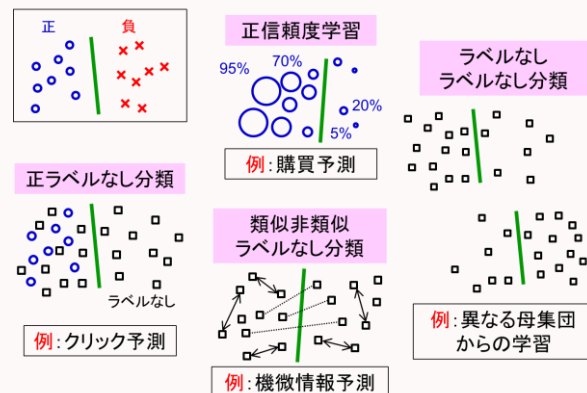
概要

機械学習では、一般に教師情報のついた訓練データを大量に必要とするが、多くの実問題では教師情報の収集が困難。容易に集められる弱い教師情報だけからでも学習を可能にする新しい弱教師付き学習理論を世界に先駆けて確立。英語専門書をMIT Pressより出版。



様々な弱教師付き学習の例

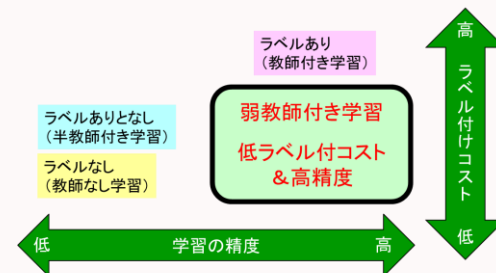
弱教師付き分類(2クラス)



表彰等

- 令和4年度文部科学大臣表彰科学技術賞研究部門
- ICLR2022 Outstanding Paper Honorable Mention (3391編中トップ0.3%)
- 音響応用でICASSP2023 Best Paper Award (6127編中1編)
- ICLR2023にて招待講演

教師付き学習からのパラダイムシフト



■ 様々な弱教師付きデータをフル活用するための理論体系を確立
Sugiyama, Bao, Ishida, Lu, Sakai & Niu. Machine Learning from Weak Supervision, MIT Press, 2022.

代表的な研究成果(ロバストベイズ深層学習)



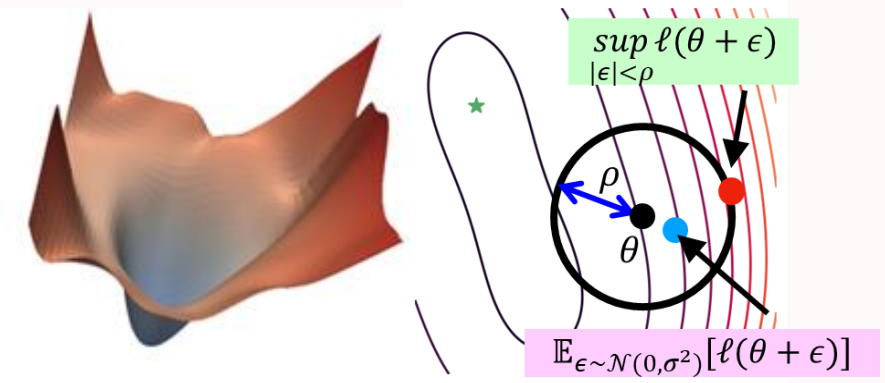
深層学習の汎化性能を向上させる実用的なアルゴリズムを開発

成果発表 ICLR 2023 (オーラル発表)

Mohammad Emtiyaz Khanら

概要

深層学習では、「平らな」局所解を見つけた方が汎化するという理論・実験的知見に基づいた Sharpness-Aware Minimization (SAM) と呼ばれる学習アルゴリズムがある。SAMがベイズ推論のFenchel双共役を用いた近似手法とみなせることを理論的に証明し、SAMの汎化性能を更に向上させる実用的なアルゴリズムを開発した。



図：平らな局所解

SAMは最大値、ベイズ推論は平均値をとる。

成果の意義・社会貢献への期待

本成果は、SAMの理論的な性質を解明するとともに、汎化性能を更に向上させることに成功した画期的な成果である。大規模言語モデルなどの基盤モデルの効率的な学習にも活用できる汎用的な技術であり、今後の実用化が大きく期待できる。

代表的な研究成果（多レベル最適化理論）



武田 朗子ら

多レベル最適化の理論とアルゴリズム

成果発表 2021年9月JMLR、2021年12月NeurIPS2021

概要

一部の変数が別の最適化問題の最適解となっている最適化問題を、多レベル最適化と呼ぶ。機械学習モデルのハイパーパラメータ最適化や、学習システムの敵対的攻撃からの防御は、多レベル最適化問題として定式化される。本成果では、微分不可能な2レベル最適化問題と微分可能多レベル最適化問題に対して、最適性保証付きの解法を初めて与えた。

$$\begin{aligned} \min_{x, y, z} f(x, y, z) \\ \text{s.t. } y \in \arg \min_{y', z'} g(x, y', z') \\ \text{s.t. } z' \in \arg \min_{z''} h(x, y', z'') \end{aligned}$$

図1：3レベル最適化問題

成果の意義・社会貢献への期待

数十年におよぶ多レベル最適化の未解決問題を肯定的に解決した画期的なものであり、学術的な意義が大きい。学習システムの信頼性向上に資すると期待。

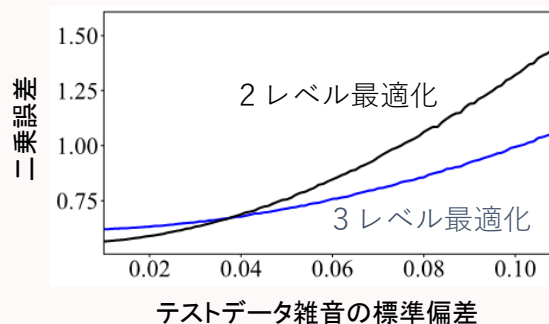


図2：機械学習のハイパーパラメータ最適化への応用

代表的な研究成果（因果推論）



清水 昌平ら

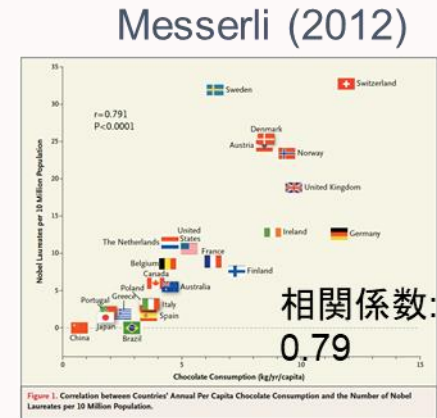
隠れ要因の存在下でも，全体構造が推定可能な因果推論の手法を初めて開発

成果発表 AISTATS (2020), UAI (2021)

概要

機械学習はデータの相関を元に予測を行うが、相関と因果は異なるため、因果関係を明らかにする因果推論が重要である。因果推論の研究においては隠れ要因の存在が最大の課題の一つであったが、本成果では隠れ要因の存在下でも、全体構造が推定可能な初めての手法を開発。

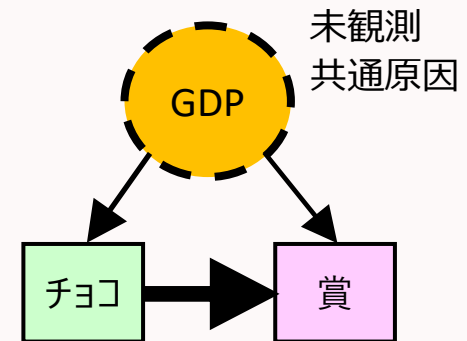
ノーベル賞受賞者数



チョコレート消費量

成果の意義・社会貢献への期待

因果推論と一見関係のない、音声分離技術の技術を活用した、隠れ要因を抽出する独創的な新手法。公開したソフトは、国内外の企業が利用を開始。



研究コミュニティでの国際的な活躍

国際学会のプログラム委員長や招待講演を行うなど
研究コミュニティでの活動度・認知度の上昇

■ General Chair:

- NeurIPS2016 (杉山)
- ACML2017 (杉山)

■ Program Chair:

- AISTATS2019 (杉山)
- ACML2019 (鈴木)
- ACML2020 (杉山)
- ACML2022 (Khan)
- IEEE-CAI2024 (杉山)
- ICLR2024 (Khan)

■ 受賞:

- ICASSP2023 Best Paper Award
- ICLR 2022 Outstanding Paper Honorable Mention
- ICLR 2021 Outstanding Paper Award
- IEEE Systems, Man, and Cybernetics 2020 Awards
- AAI2020 Outstanding Paper Award Honorable Mention
- ACML 2019 Best Paper Award
- IEEE-ICASSP2019 Best Student Paper Award

■ Tutorial:

- CoRL2018 (杉山)
- NeurIPS2019 (Khan)
- QTML2019 (Minh)
- ACML2020 (Zhao)
- SPCOM2020 (Khan)
- ACML2021 (杉山, Khan, 鈴木)
- IJCAI2021 (杉山)
- CIKM2022 (杉山)
- ACPR2021 (鈴木)

■ Keynote:

- ALT-DS2017 (杉山)
- CCAI2017 (杉山)
- IWAENC2018 (杉山)
- ICONIP2018 (杉山)
- CCF-ICAI2018 (杉山)
- ICDM2021 (杉山)
- ICCOPT2022 (武田)
- MaxEnt2022 (Khan)
- ACML2022 (鈴木)
- ALT2023 (鈴木)
- ICLR2023 (杉山)

今後の展望

■ 大規模言語モデル・生成AIの時代の課題:

● トランスフォーマー・拡散モデルの理論:

- 収束率
- 最適化ランドスケープ

● 大規模な基盤モデルの学習:

- 弱い教師情報、ノイズの多い教師情報からの学習
- 偏りのあるデータや敵対的なデータからの学習
- モデルやデータ圧縮

● 予測と推論を超えて:

- 因果推論に基づく介入
- 逐次的意思決定

■ 国内外の出口研究組織との連携を更に強化

■ 人材不足に対応すべく雇用の安定化・国際的な人材登用を加速