文科省・情報科学技術分野における
戦略的重要研究開発領域に関する検討会

資料9
科学技術・学術審議会　情報委員会
情報科学技術分野における戦略的重要研究
開発領域に関する検討会（第1回）
令和6年4月24日

2024年4月24日

AIのサイエンス

杉山　将

理化学研究所　革新知能統合研究センター
／東京大学

http://www.ms.k.u-tokyo.ac.jp/sugi/

自己紹介

現職：

- 理化学研究所・センター長：研究者とともに
- 東京大学・教授：学生とともに
- 企業・技術顧問：経営者，エンジニアとともに

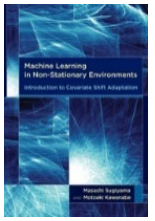
専門分野：

- 機械学習の理論とアルゴリズム開発
- 英語著書→

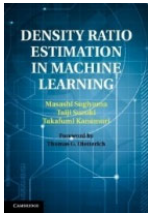
学会活動：

- NeurIPS2015, AISTATS2019, ACML2010/2020 プログラム委員長
- ICLR2023キーノート講演
- 通信学会IBISML研究会委員長(2022-24)

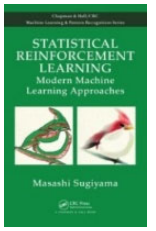
Sugiyama & Kawanabe, Machine Learning in Non-Stationary Environments, MIT Press, 2012



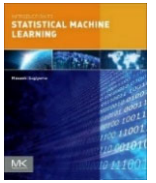
Sugiyama, Suzuki & Kanamori, Density Ratio Estimation in Machine Learning, Cambridge University Press, 2012



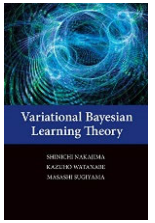
Sugiyama, Statistical Reinforcement Learning, Chapman and Hall/CRC, 2015



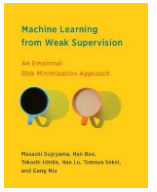
Sugiyama, Introduction to Statistical Machine Learning, Morgan Kaufmann, 2015



Nakajima, Watanabe & Sugiyama, Variational Bayesian Learning Theory, Cambridge University Press, 2019



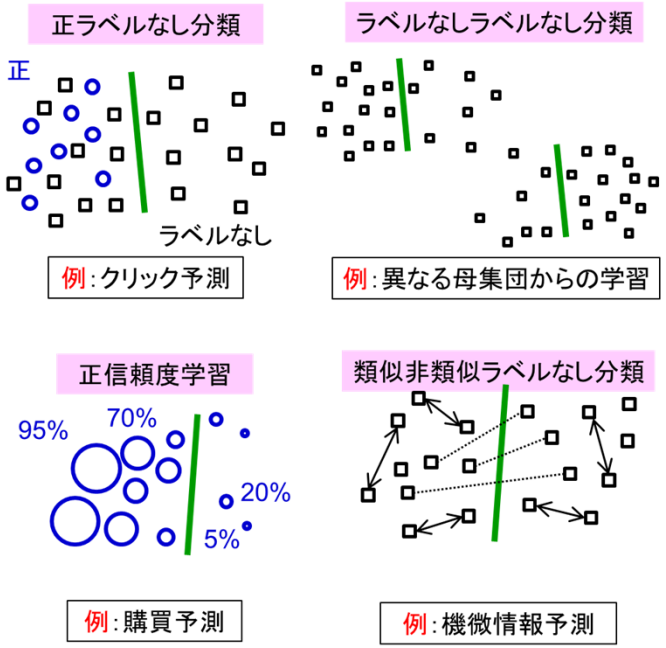
Sugiyama, Bao, Ishida, Lu, Sakai & Niu. Machine Learning from Weak Supervision, MIT Press, 2022.



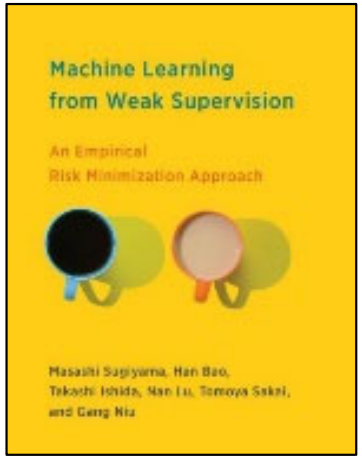
最近の研究：弱教師付き学習

- 収集コストの低い
  「弱い」教師データ
  だけから学習

Wang, Xiao, Li, Feng, Niu, Chen & Zhao
(ICLR2022 Outstanding Paper Honorable Mention),
Sugiyama (ICLR2023 Plenary Talk)



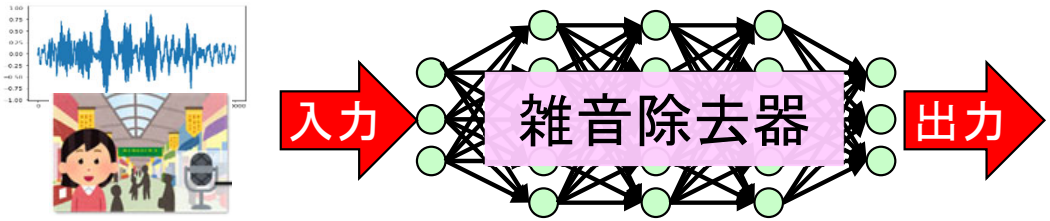
Sugiyama, Bao, Ishida,
Lu, Sakai & Niu.
(MIT Press 2022)

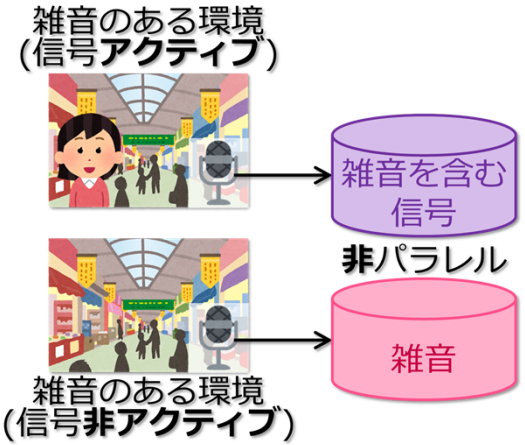


- 例：音声の雑音除去

Ito & Sugiyama
(ICASSP2023 Best Paper Award)

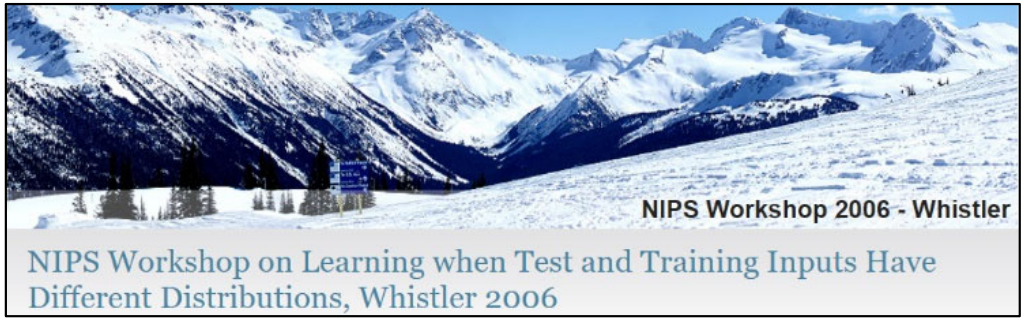
- 雑音を含む音声と雑音だけから学習
  できる！（収集コストのかかる
  雑音を含まない音声は不要）

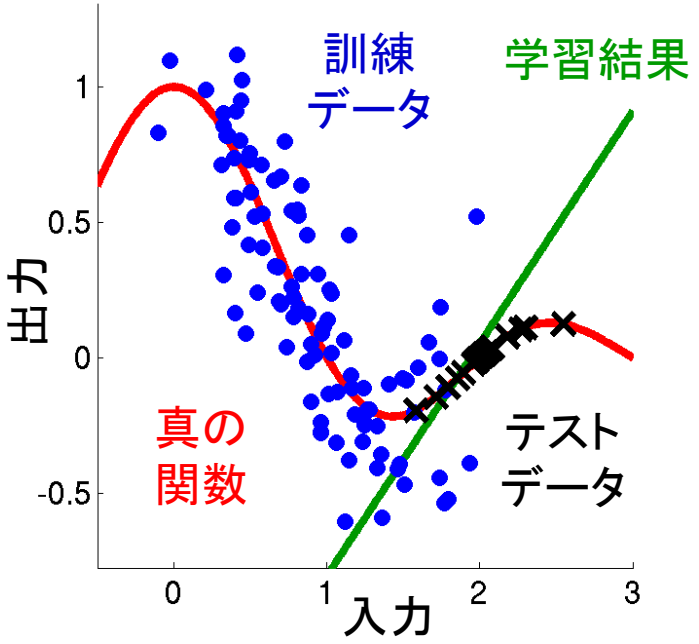




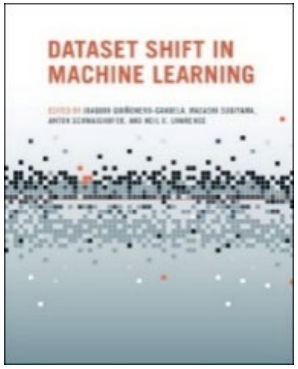
最近の研究：転移学習

- 訓練データとテストデータの分布の違いを吸収
  - 古典的な設定：共変量シフト（入力分布だけが変化） Shimodaira (JSPI2000)





Quiñonero-Candela, Sugiyama, Schwaighofer & Lawrence (MIT Press 2009)

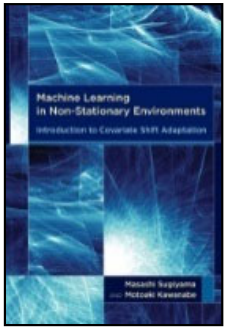


  - 重要度重み付き学習:

$$\underset{f\in\mathcal{F}}{\operatorname{argmin}} \left[\sum_{i=1}^{n_{\mathrm{tr}}} \frac{p_{\mathrm{te}}(\boldsymbol{x}_i^{\mathrm{tr}})}{p_{\mathrm{tr}}(\boldsymbol{x}_i^{\mathrm{tr}})} \ell(f(\boldsymbol{x}_i^{\mathrm{tr}}), y_i^{\mathrm{tr}}) \right]$$

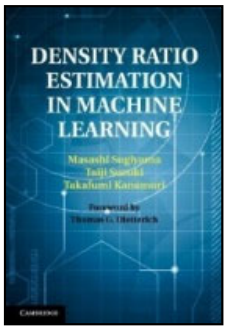
重要度　　損失

重要度（密度比）の推定が重要！



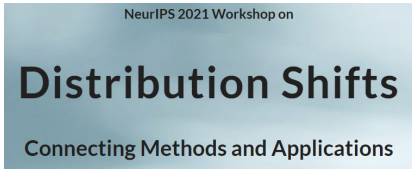
Sugiyama & Kawanabe (MIT Press 2012)

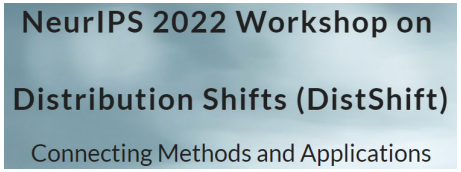
Sugiyama, Suzuki & Kanamori (Cambridge University Press 2012)



最近の研究：転移学習

- 近年再注目：NeurIPS2021, 2022, 2023ワークショップ他





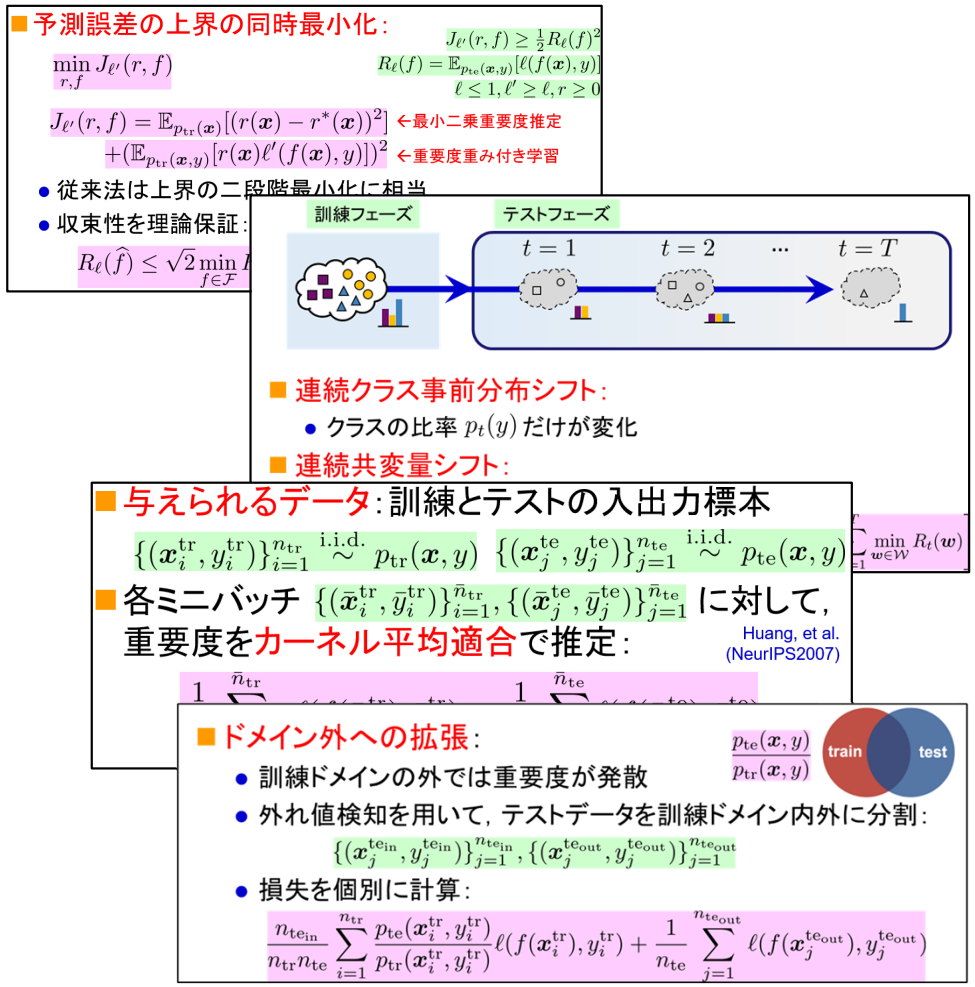
- 最近の発展：
  - 重みと予測器の同時学習

    Zhang, Yaname, Lu & Sugiyama
    (ACML2020 Best Paper Award, SNCS2021)
  - 連続分布シフトへの拡張

    Bai, Zhang, Zhao, Sugiyama & Zhou (NeurIPS2022)
    Zhang, Zhang, Zhao & Sugiyama (NeurIPS2023)
  - 同時分布シフトへの拡張

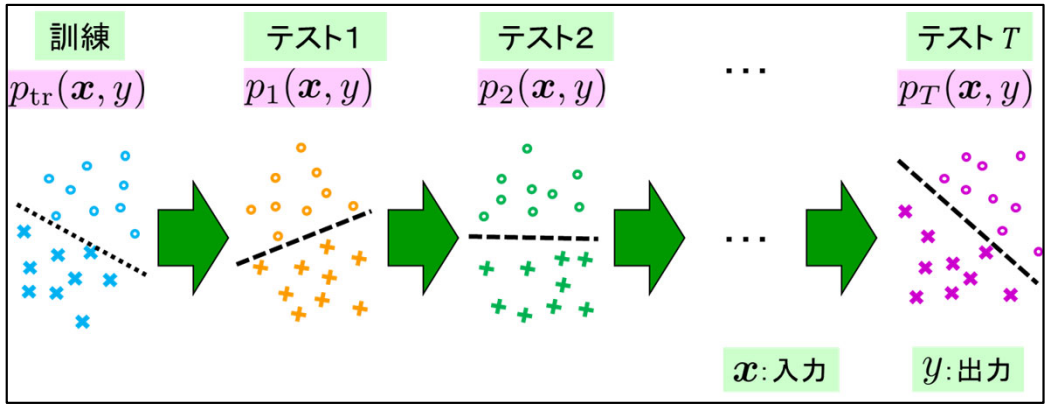
    Fang, Lu, Niu & Sugiyama
    (NeurIPS2020 Spotlight)
  - 分布外適応への拡張

    Fang, Lu, Niu & Sugiyama
    (NeurIPS2023 Spotlight)



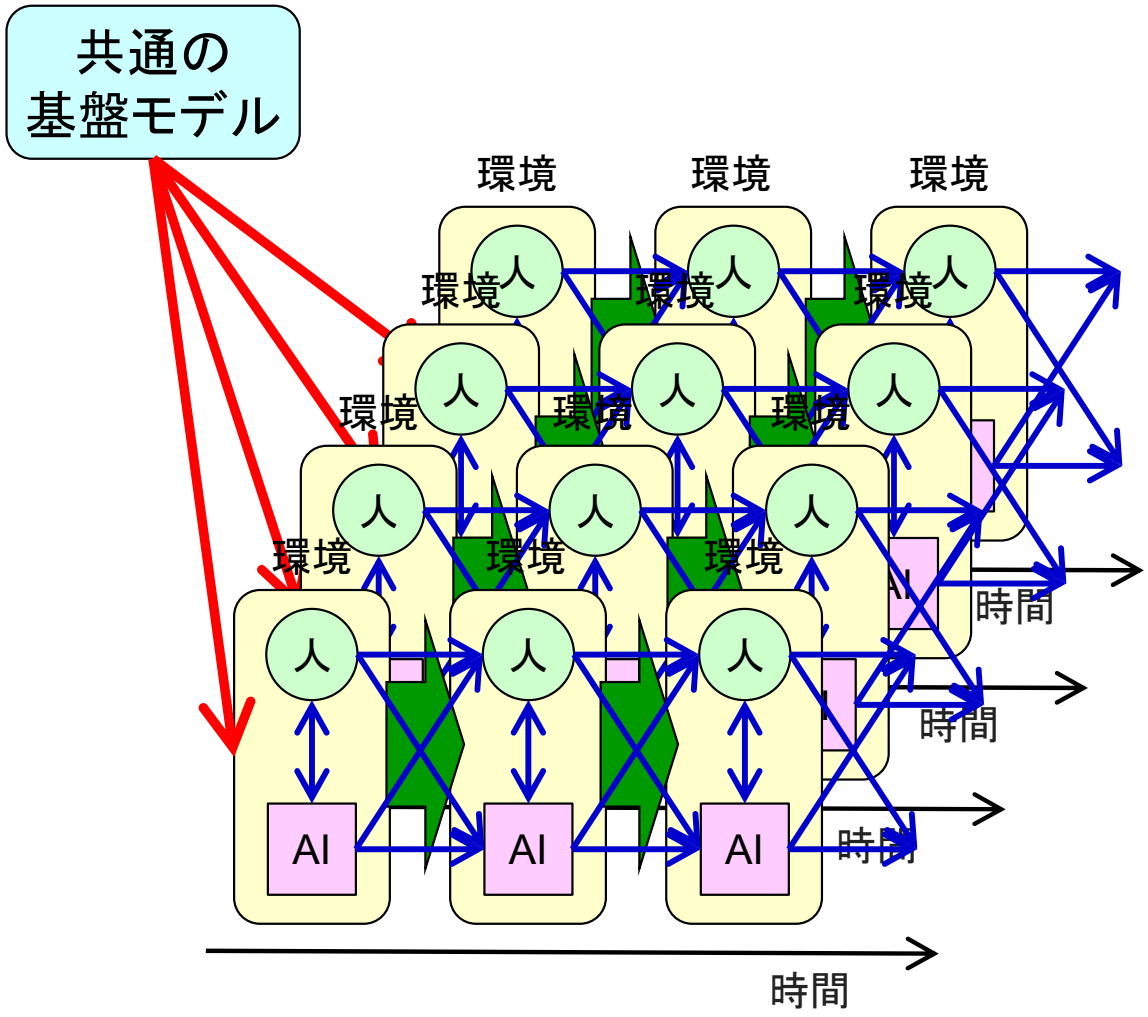
今後の課題：無仮定機械学習

- これまで：
  - （共変量シフトなど）適切な仮定のもとでの理論とアルゴリズム
  - 仮定が正しければうまくいく．そうでなければ保証なし
- 現実：
  - 仮定が成り立っているか分からない
  - やってみてうまくいったらラッキー
- 目標：
  - （ほぼ）無仮定機械学習の理論とアルゴリズム開発
  - （ほぼ）任意の場面で最低限の性能が保証される
  - 現場で最初に試すべき極限的な汎用手法！
  - 例：弱教師付き連続同時分布シフト適応



今後の課題：個別化AIの全体最適化

- 中央にある巨大な共通の基盤モデルを，個人ごと，タスクごとに個別化：
  - 各AIがユーザと相互作用しながら時間発展
  - AI同士，ユーザ同士も相互作用
- 社会全体で最適化：
  - 効率性
  - 信頼性
  - 説明性
  - 公平性
  - 安全性



今後の課題：個別化AIの全体最適化

- **学習の信頼性向上**：
  - 不完全情報からの弱教師付き学習
  - 低信頼性データからのロバスト学習
  - 予測の不確実性に基づく棄却
  - データの異常・変化検知
  - データ改変に対する敵対的学習
  - 国，文化，人種等の多様性のドメイン適応
  - 非定常環境，タスク変化に対する継続学習
  - 言語，画像，音響，科学データ等のマルチモーダル学習
- **理論解析**：
  - 深層学習の汎化解析
  - 非凸・確率的・分散最適化の収束解析
  - 学習理論と最適化理論の融合

AI開発の現在地 ㊦

杉山将 東京大学教授

原理解明 さらなる研究必要

日経新聞　経済教室（2023年3月30日）

画像生成やテキスト対話型の人工知能（AI）システムが注目を集めている。専門的な知識は必要とせず、簡単な言葉を入力するだけで実用的な出力が得られるため、AIの活用は新たな局面に突入したといえそうだ。本稿では、AIシステムの基本となる機械学習の技術的な概要と、その本質について述べたい。

機械学習とは、機械（コンピューター）に学習能力を持たせるための情報技術の総称である。音声や画像の認識、言語の翻訳、情報の検索、商品の推薦、顧客情報の分析、株価の予測、サイバー攻撃の検知、自動運転車の制御など、様々な場面で用いられている。

メディアで「AIが学習する」という言葉を頻繁に見聞きする。擬人化されたAIが私たちの思いをくみ取って自発的に学習してくれるわけではなく、技術者がプログラムをつくってコンピューターを学習させて

すぎやま・まさし
74年生まれ。東京工業大博士（工業）。専門は機械学習。理化学研究所センター長

- **大規模モデル・データ対応**：
  - 個別化モデルのマルチタスク学習
  - モデルを逐次更新するオンライン学習
  - 不確定性を考慮した近似ベイズ推論
  - モデルを小型化するモデル圧縮
  - 圧縮データからの直接学習
  - 分散データからの連合学習
  - 巨大モデルの分散最適化
- **新しい推論機構**：
  - データからの因果推論
  - 基盤モデルを用いた文脈内学習
  - 長期的な予測に基づく逐次的意思決定
  - 動的システムに基づく時系列予測
  - データ駆動型仮説検定
- **社会的観点**：
  - 予測・モデルの説明性
  - プライバシ保護
  - 意思決定・資源配分の公平性
  - 多様なAIと人間の相互作用
  - モデル・ユーザ間通信の安全性

AI研究人材

- 日本の国際的な研究競争力は限定的:
  - 日本人の博士進学者数は増加せず
  - 博士支援を改善しても，タイパの悪い研究はしたくない
  - 博士学生不足が研究者・教員不足に直結
  - 少子高齢化の加速で，今後も抜本的な改善は見込めない
  - 頼みの留学生も，少子化・経済安保の影響を受ける見込み
  - 社会人の再教育は，必ずしも高度研究人材育成につながらない
  - 女子学生の情報系進学の支援は，効果は低いが絶対必要
  - （親御さん・小中高の進路担当教員・予備校への働きかけ）
- 世界の優秀な若手人材はスター研究者の周りに結集:
  - 安定してボトムアップに研究ができる環境を整え，
    日本発の国際的スター研究者を育成すべき
  - 海外人材のクロアポ雇用できる仕組みの導入