

生成AIがもたらす社会変化と 今後必要になる取組

理化学研究所／東京大学

杉山 将



自己紹介

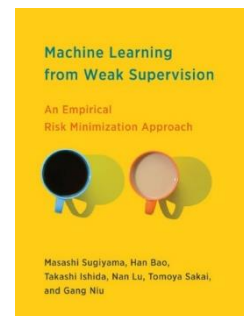
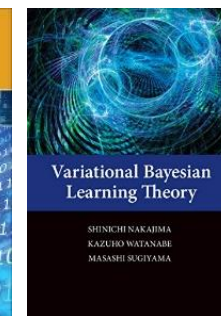
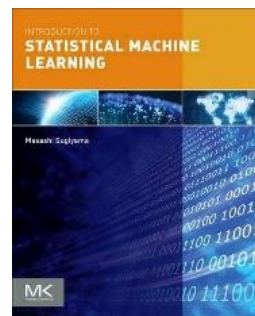
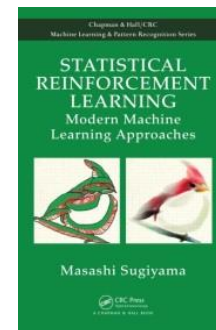
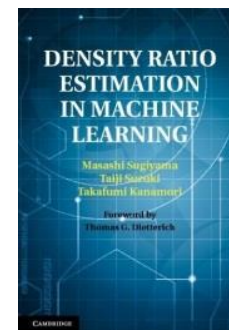
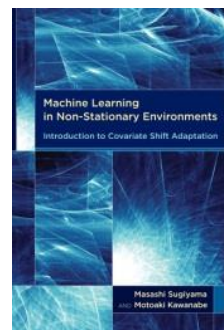
■ 現職:

- 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:

- 機械学習の数学的な基礎研究
(コンピュータ, 統計学など)
- 機械学習技術の実世界応用
(画像, 言語, 音声, 脳波,
ロボット, 医療, 生命など)



■ 研究体制:

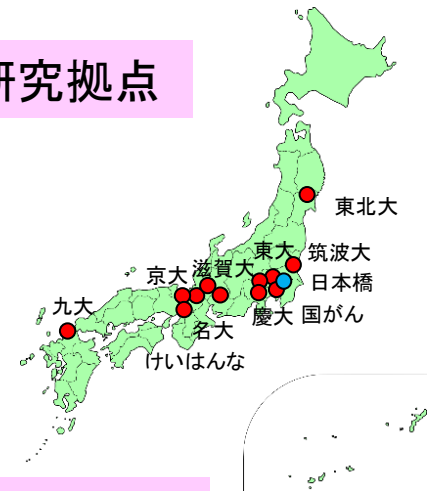
- 汎用基盤技術研究グループ(杉山)
- 目的指向基盤技術研究グループ(上田)
- 社会における人工知能研究グループ(橋田)



■ 統計情報:

- 常勤研究員128名(外国人41%, 女性23%)
- 客員研究員237名, 学生100名
- 延べ146名の海外インターン生
- 40+企業と共同研究
- 40+海外大学・研究機関と連携

研究拠点



日本橋オフィス



AIPの3つの研究グループ

汎用基盤技術 (12チーム)

不完全情報学習チーム チームリーダー 杉山 将 (Ph.D.) 研究室WEB 理研WEB	構造的学習チーム チームリーダー 河原 吉伸 (Ph.D.) 研究室WEB 理研WEB
テンソル学習チーム チームリーダー Qibin Zhao (Ph.D.) 研究室WEB 理研WEB	関数解析的学習チーム チームリーダー Quang Minh Ha (Ph.D.) 研究室WEB 理研WEB
圧縮情報処理ユニット ユニットリーダー 田部井 靖生 (Ph.D.) 研究室WEB 理研WEB	深層学習理論チーム チームリーダー 鈴木 大慧 (Ph.D.) 理研WEB
計算論的学習理論チーム チームリーダー 畑笠 晃平 (Ph.D.) 研究室WEB 理研WEB	因果推論チーム チームリーダー 清水 昌平 (Ph.D.) 研究室WEB 理研WEB
近似ベイズ推論チーム チームリーダー Mohammad Emteyaz Khan (Ph.D.) 研究室WEB 理研WEB	連続最適化チーム チームリーダー 武田 朗子 (Ph.D.) 研究室WEB 理研WEB
数理科学チーム チームリーダー 坂内 健一 (D.Math.Sci.) 研究室WEB 理研WEB	高次元因果解析チーム チームリーダー 今泉 允聡 (Ph.D.) 理研WEB

- 学習と推論のアルゴリズム
 - ロバスト, ベイズ, 構造, ...
- 統計的学習と最適化の理論
 - 収束性, 表現力, 非凸, ...

目的指向基盤技術 (17チーム)

がん探索医療研究チーム チームリーダー 浜本 隆二 (Ph.D.) 理研WEB	IPS細胞連体医学的リスク回避チーム チームリーダー 上田 修功 (Ph.D.) 理研WEB
分子情報科学チーム チームリーダー 津田 宏治 (Ph.D.) 理研WEB	認知行動支援技術チーム チームリーダー 大武 美保子 (Ph.D.) 研究室WEB 理研WEB
防災科学チーム チームリーダー 上田 修功 (Ph.D.) 研究室WEB 理研WEB	インフラ管理ロボット技術チーム チームリーダー 岡谷 貴之 (Ph.D.) 理研WEB
医用機械知能チーム チームリーダー 原田 達也 (Ph.D.) 理研WEB	データ駆動型生物医学科学チーム チームリーダー 竹内 一郎 (Ph.D.) 研究室WEB 理研WEB
計算脳ダイナミクスチーム チームリーダー 山下 直人 (Ph.D.) 研究室WEB 理研WEB	遺伝統計学チーム チームリーダー 田宮 元 (Ph.D.) 研究室WEB 理研WEB
病理情報学チーム チームリーダー 山本 龍一朗 (M.D., Ph.D.) 理研WEB	自然言語理解チーム チームリーダー 乾 健太郎 (Ph.D.) 研究室WEB 理研WEB
知識獲得チーム チームリーダー 松本 祐治 (Ph.D.) 理研WEB	言語情報アクセス技術チーム チームリーダー 関根 聡 (Ph.D.) 理研WEB
空間情報学チーム チームリーダー 横矢 進人 (Ph.D.) 研究室WEB 理研WEB	音楽情報知能チーム チームリーダー 浜中 雅俊 (Ph.D.) 研究室WEB 理研WEB
音響情報理解チーム チームリーダー 吉井 和佳 (Ph.D.) 研究室WEB 理研WEB	

社会におけるAI (6チーム)

社会におけるAI利活用と法制度チーム チームリーダー 中川 裕志 (Ph.D.) 理研WEB	科学技術と社会チーム チームリーダー 佐倉 統 (Ph.D.) 理研WEB
分散型ビッグデータチーム チームリーダー 楡田 浩一 (Ph.D.) 理研WEB	経済経営情報融合分析チーム チームリーダー 星野 崇宏 (Ph.D.) 研究室WEB 理研WEB
人工知能セキュリティ・プライバシーチーム チームリーダー 佐久間 淳 (Ph.D.) 研究室WEB 理研WEB	人工知能安全性・信頼性ユニット ユニットリーダー 荒井ひろみ (Ph.D.) 理研WEB

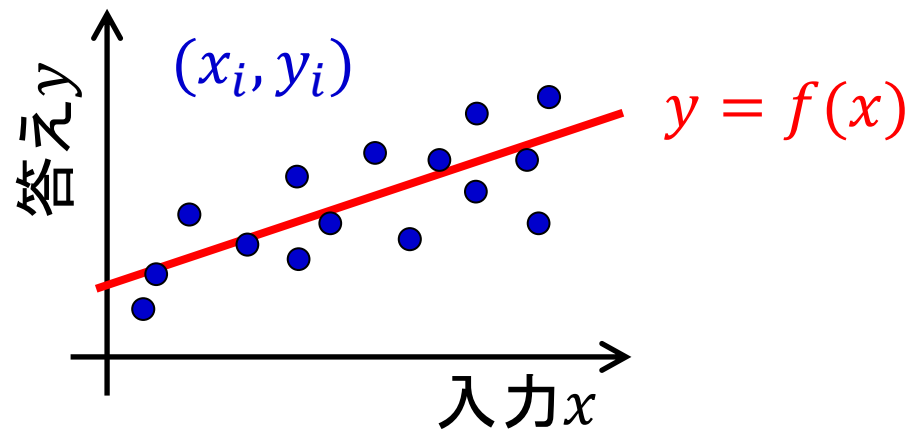
- 個人データ管理
- AIセキュリティ
- 経済と経営
- 倫理とガバナンス

- AIによる科学研究の加速
 - がん, 材料, 遺伝子, ...
- AIによる社会課題解決への貢献
 - 自然災害, 高齢者ヘルスケア, 教育, ...

通常の機械学習

■ 識別モデルの教師付き学習:

- 質問 x に対する答え y を教え、
入出力関係 $y = f(x)$ を学習



- 最小二乗法: $\min_g \sum_i (y_i - g(x_i))^2$
- 未知の入力 x' に対する答え y' が予測できる

識別モデルの例

音声認識(声→言語)



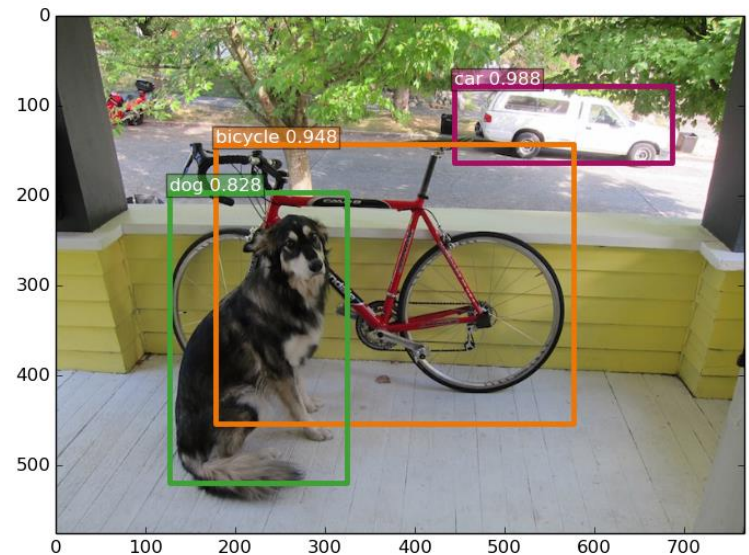
<https://spjai.com/speech-recognition/>

言語翻訳(日本語→英語)



<https://toyokeizai.net/articles/-/167690>

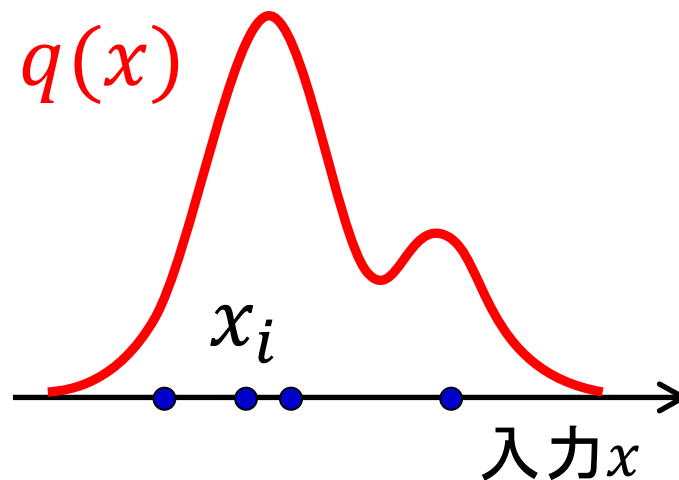
画像理解(画像→物体)



<https://github.com/zhreshold/mxnet-ssd>

■ 生成モデルの教師なし学習:

- データ x の例から, データの生成確率 $p(x)$ を学習



- 最尤推定法: $\max_q \prod_i q(x_i)$
- 学習した確率分布もとに, 新しいデータ x' が生成できる

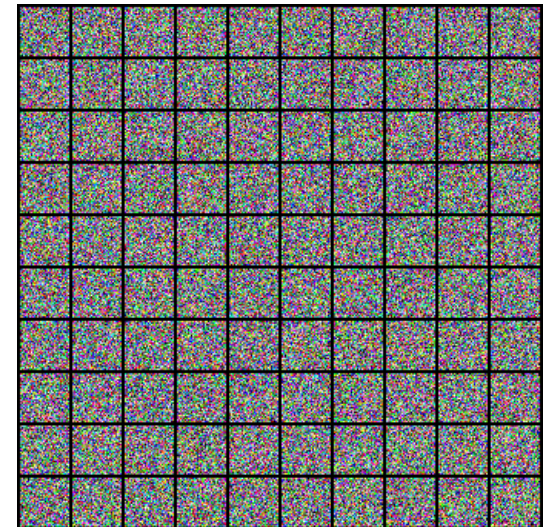
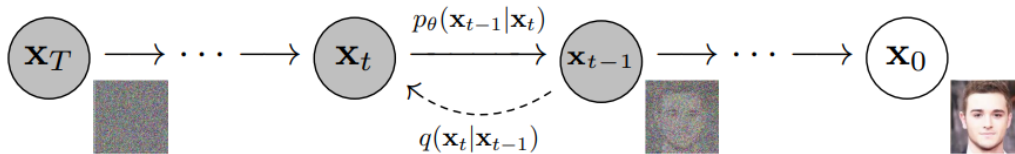
生成モデルの例：画像

■ 拡散モデル

- **学習**：画像に徐々に雑音を加えていく
- **生成**：雑音から徐々に雑音を除去



<https://arxiv.org/pdf/2006.11239.pdf>



<https://yang-song.net/blog/2021/score/>

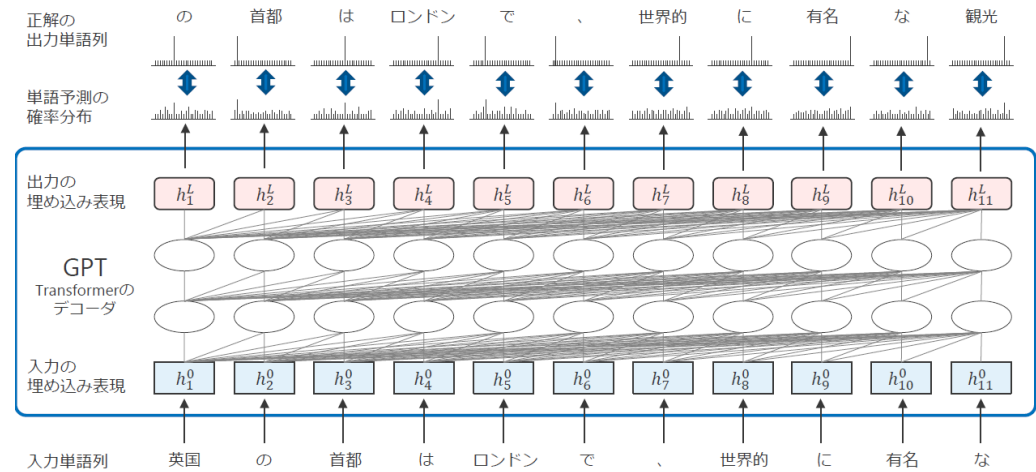
生成モデルの例：言語

■ GPT (Generative Pre-trained Transformer)

- **学習**：文章の分布を学習
- **生成**：与えられた文章の次の単語を生成

ChatGPT		
☀ Examples	⚡ Capabilities	⚠ Limitations
"Explain quantum computing in simple terms" →	Remembers what user said earlier in the conversation	May occasionally generate incorrect information
"Got any creative ideas for a 10 year old's birthday?" →	Allows user to provide follow-up corrections	May occasionally produce harmful instructions or biased content
"How do I make an HTTP request in Javascript?" →	Trained to decline inappropriate requests	Limited knowledge of world and events after 2021

<https://chat.openai.com/>



https://speakerdeck.com/chokkan/20230327_riken_llm

■ 日本独自の生成AIは必要か？

- Yes. 仮に技術的に新規性がなくとも、日本の文化を反映させた生成AIを普及させるべき

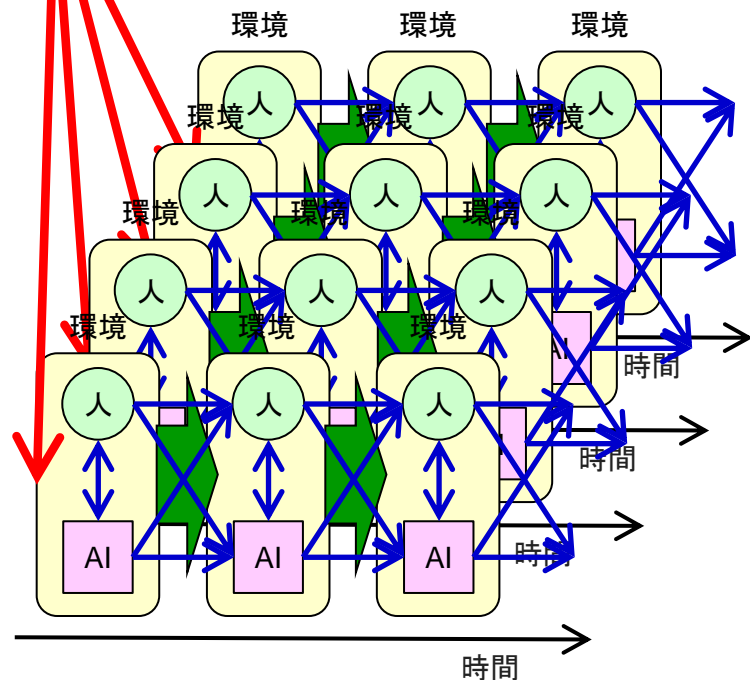
■ 今の言語モデルは、事後確率最大の解を出力？

- 多様な意見があるはずなので、事後分布を示すべき

■ そもそも、1つの言語モデルを皆で使うのはおかしい？

- 人に個性があるように、言語モデルも個別化すべき

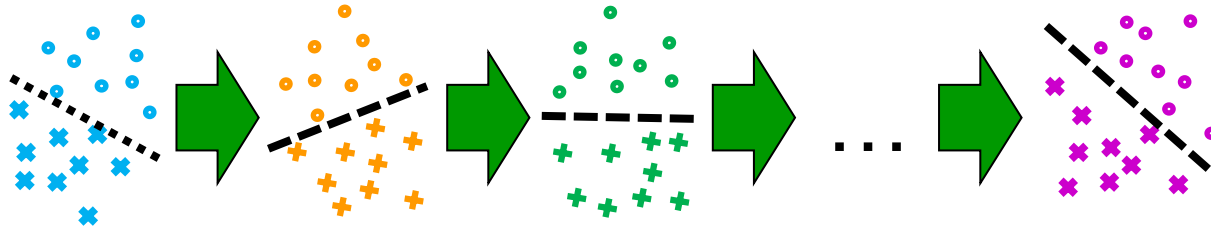
共通の基盤モデル



重要な研究課題

■ オンライン学習:

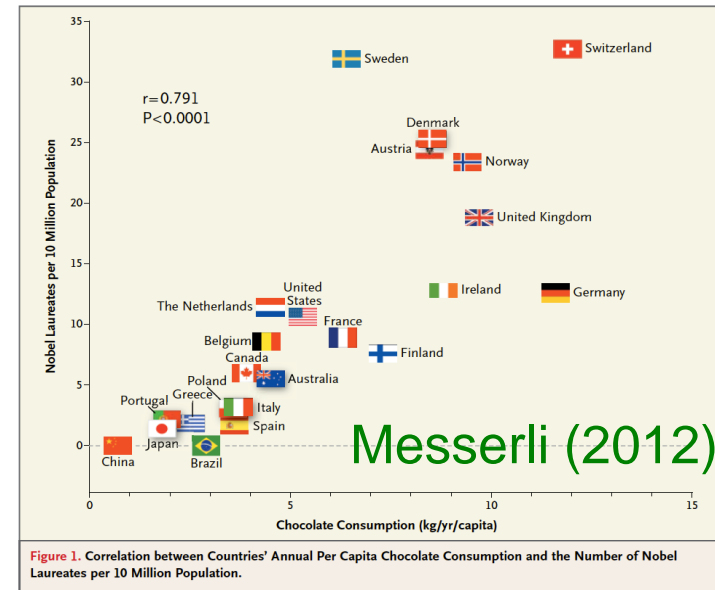
- 基盤モデルに最新の情報を追加していく必要がある
- 巨大モデルの再学習は不可能. 逐次学習が不可欠



■ 因果推論:

- 相関関係だけでなく, 因果関係の分析が重要
- 介入した場合の確率の変化や, 反実仮想的な確率を推測

ノーベル賞受賞者数

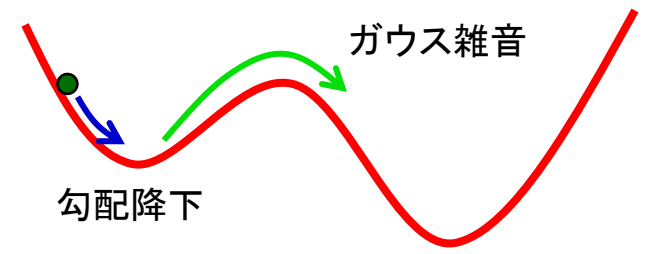
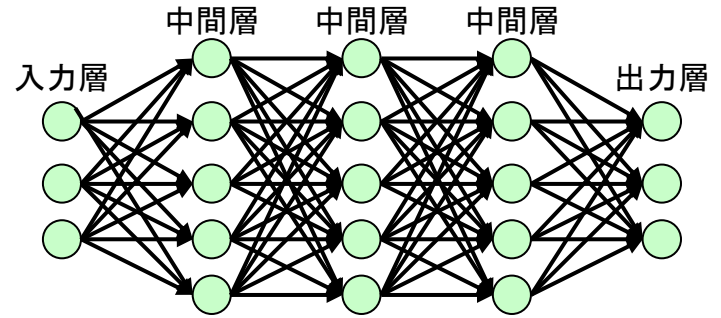


チョコレート消費量

重要な研究課題

■ 深層学習の理論:

- 深層学習はなぜ浅層学習より予測性能が高いのか？
- トランスフォーマーの汎化能力は？
- なぜ大域的最適化が可能なのか？



■ 弱教師付き学習:

- 限られた教師情報からの高精度学習

■ 雑音ロバスト学習:

- 教師情報に含まれる雑音を抑制しながら高精度に学習

弱教師付き学習

正ラベルなし分類

例: クリック予測

du Plessis et al. (NIPS2014, ICML2015, MLJ2017), Niu et al. (NIPS2016), Kiryo et al. (NIPS2017), Hsieh et al. (ICML2019)

正信頼度学習

例: 購買予測

Ishida et al. (NeurIPS2018), Shinoda et al. (IJCAI2021)

ラベルなしラベルなし分類

例: 異なる母集団からの学習

du Plessis et al. (TAAD2013), Lu et al. (ICLR2019, AISTATS2020), Charoenphakdee et al. (ICML2019), Lei et al. (ICML2021)

類似非類似ラベルなし分類

例: 機械情報予測

Bao et al. (ICML2018), Shimada et al. (NeCo2021), Dan et al. (ECMLPKDD2021), Cao et al. (ICML2021), Feng et al. (ICML2021)

ラベル雑音に対するロバスト学習理論

背景:

- ラベル雑音が存在するとき、大量の訓練データを用いても学習はうまくいかない
- 雑音を除去する明示的な機構が必要

従来のアプローチ:

- 教師なし異常値除去: 分類よりも難しく信頼性が低い
- ロバスト損失, 正則化による平滑化: 回帰(加法雑音)には有効だが、分類(ラベル変化)には不十分

一連の成果: ラベル雑音学習の理論体系を構築

- ラベルの変化確率(雑音遷移行列)の明示的なモデル化
- 雑音の無いデータがあれば、雑音遷移行列は推定可能
- 雑音のあるデータだけでは、数学的には同定不可能
- 雑音遷移行列と分類器の同時学習
- 入力依存ラベル雑音への拡張

1	0.1	0.5
0	0.8	0.5
0	0.1	0

Berthon, Han, Niu, Liu & Sugiyama, ICML2021 Long
Zhang, Niu & Sugiyama, ICML2021 Long
Li, Liu, Han, Niu & Sugiyama, ICML2021
Yao, Liu, Gong, Han, Niu & Zhang, NeurIPS2021

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美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る。ただ、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)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」という言葉を頻りに見聞きする。擬人化されたAIが私たちの思いをくみ取って自動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うまくコンピュータを学習させて



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

A I 開発の現在地 ①

杉山将 東京大学教授

原理解明さらなる研究必要

<https://www.nikkei.com/article/DGXZQOCD152XM0V10C23A3000000/>

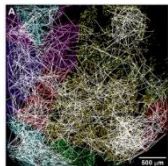
- さて、このような新AI時代にどう対応すべきだろうか。少なくとも、人間によるAI技術を活用する能力と、**AIが生成した情報の真贋を見抜く力**は必要だ。
- ただし**真贋の判定**はまさに機械学習における**教師付き学習**問題そのものであり、人間が真贋を判断するとは、結局はAIが自動生成した入力データに、**人間が教師情報を付与するのと同じこと**である。
- いまはAI技術の現実的な発展のために**身を粉にして教師情報作りを続ける**べきか、それとも新たな知性の獲得のため**原理の解明を目指す**べきか。まさに議論が必要である。

理化学研究所における生成AI研究¹⁴

DXを活用し、異分野の連携融合を促進する「つなぐ科学」

先端計算科学と予測アルゴリズム、データ整備を連環させ、
未来の予測制御の科学を開拓

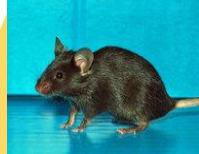
良質なデータ整備



脳の神経活動のリアルタイム可視化技術(例)



SPring-8 SAGLA



バイオリソース

研究DXの先駆的取組へ発展

データ

AI

スパコン

基盤モデル・生成AI

計算可能領域の拡張
量子古典ハイブリット

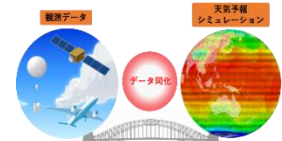


量子コンピュータ

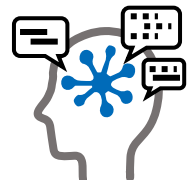


スパコン「富岳」

AI×数理で
予測の科学を開拓



学習理論



AI

狙い
研究サイクル
加速

「未来の予測制御の科学」を
分野の枠を超えて開拓