

AI・機械学習の最近の発展: 教師付き学習から 基盤モデル, AIエージェントへ

杉山 将

理化学研究所／東京大学

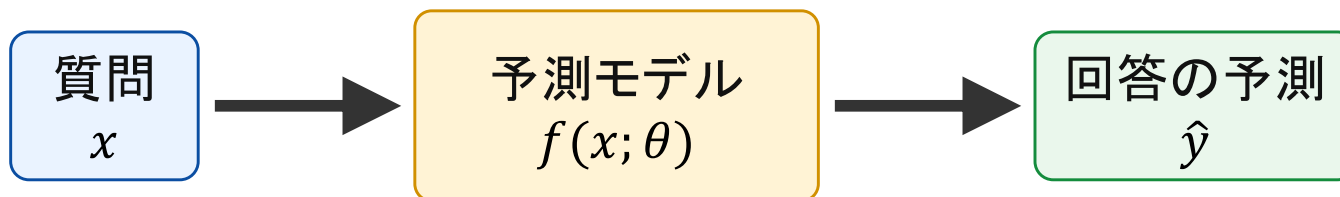


<http://www.ms.k.u-tokyo.ac.jp/sugi/>

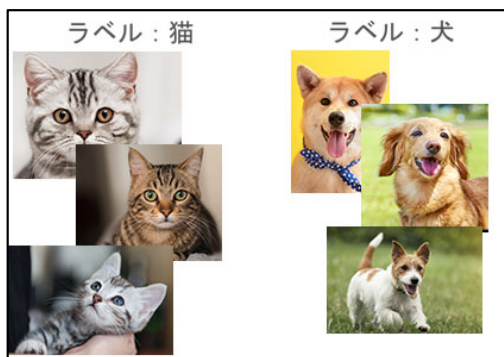


古典的な教師付き学習

- 質問に対する回答の**予測モデル**を、
質問と回答のペア(教師付き訓練データ)を用いて構築



- 例:** 画像認識(x : 画像, y : 猫/犬)



<https://www.cct-inc.co.jp/media/ai/news/primer/ai-08/>

- 迷惑メール判定
(x : メール, y : 迷惑/正常)

訓練メール x	訓練ラベル y
今だけ無料!こちらをクリック	迷惑
明日の会議は10時からです	正常
100万円が当選しました	迷惑
資料を添付しました	正常

- 問題点:**

- タスクが変わるたびに、新しいデータの収集・ラベル付けが必要
- 難しいタスクほど、大量の良質な訓練データを集めるのが困難
- モデルは基本的にそのタスク専用で、別タスクへ使い回しにくい

基盤モデルを用いた予測

■ 基盤モデル: $F(\cdot; \Theta)$

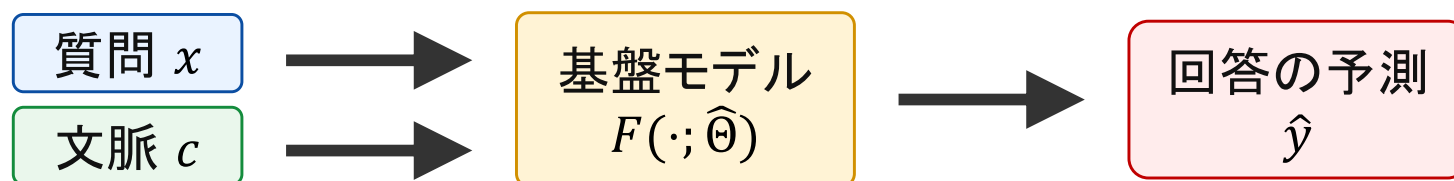
- 大量・多様なデータで学習された汎用モデル
(例: トランスフォーマ)

■ 事前学習 (汎用的な知識を学習): $\hat{\Theta}$

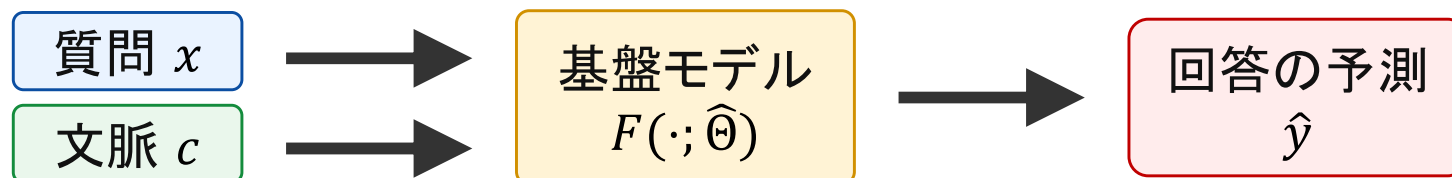
- 様々なタスクからデータを集め, 巨大なモデルを学習
(例: 系列 $x^{(1)}, \dots, x^{(T)}$ から $x^{(T+1)}$ を自己回帰学習)

■ 基盤モデルを用いた予測: $\hat{y} = F(x, c; \hat{\Theta})$

- c : 解きたいタスクの説明



基盤モデルを用いた予測



■ 例1: 迷惑メール判定

- x : 「今だけ無料！こちらをクリック」
- c : 「迷惑/正常で答えよ」
- $\hat{y}$: 「迷惑」

c に訓練データを含めることもできる
(**文脈付き学習**)

■ 例2: 算数の文章問題

- x : 「120円のリンゴを3個買い, 500円払う. お釣りは？」
- c : 「日本語で簡潔に答えよ」
- $\hat{y}$: 「お釣りは140円です」

■ **問題点**: 次の単語を予測するだけで, 人間の指示や価値観に沿った応答をするようには学習されていない

基盤モデルのさらなる改善

■ 事後学習: $\hat{\Theta} \rightarrow \hat{\Theta}'$

- 様々な指示やフィードバックを用いて、モデルを望ましいタスクや振る舞いに適応させる
(**例**: 教師付きファインチューニング, 人間のフィードバックからの強化学習)

■ テスト時計算:

- テスト質問を一度に回答するのではなく、検証・修正しながら、より良い回答を推論探索
- 深く推論／広く推論／木探索

基盤モデルのテスト時計算

6

■ 深く推論：思考連鎖など

- 段階的に深掘りし，検証・修正しながら進む

- $x \xrightarrow{F} z_1 \xrightarrow{F} z_2 \xrightarrow{F} \cdots \xrightarrow{F} z_D \xrightarrow{F} \hat{y} \left(= F(x, z_1, \dots, z_D, c; \hat{\Theta}') \right)$

■ 例：算数の文章問題

- x : 「120円のリンゴを3個買い，500円払う．お釣りは？」
- c : 「日本語で簡潔に答えよ」
- z_1 : 「代金は $120 \times 3 = 360$ 円」
- z_2 : 「お釣りは $500 - 360 = 140$ 円」
- z_3 : 「単位と計算を確認」
- $\hat{y}$: 「お釣りは140円です」

基盤モデルのテスト時計算

7

■ 広く推論: 思考木など

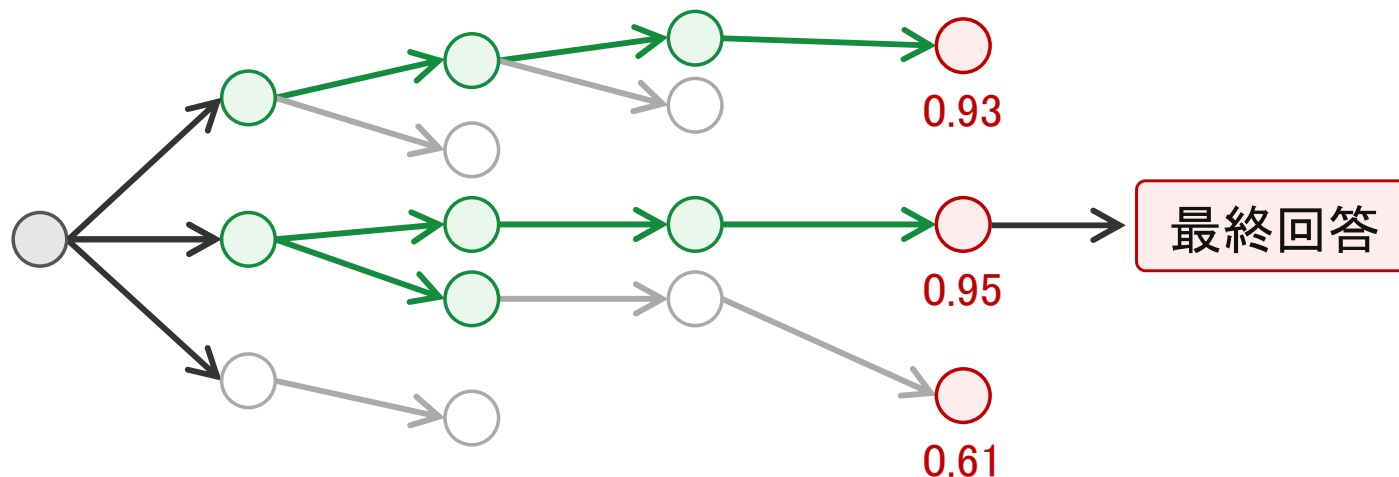
- 複数の候補経路を
並列に生成・評価し、
最も良い回答を選ぶ

$$x \xrightarrow{F} \left\{ \begin{array}{l} z_1^{(1)} \xrightarrow{F} \dots \xrightarrow{F} z_D^{(1)} \xrightarrow{F} y^{(1)} \\ z_1^{(2)} \xrightarrow{F} \dots \xrightarrow{F} z_D^{(2)} \xrightarrow{F} y^{(2)} \\ \vdots \\ z_1^{(W)} \xrightarrow{F} \dots \xrightarrow{F} z_D^{(W)} \xrightarrow{F} y^{(W)} \end{array} \right\} \xrightarrow[s(x,y)]{\text{評価・選択}} \hat{y}$$

$s(x, y)$: 外部検証器, 報酬モデル,
LLM-as-a-judge など

■ 木探索型:

- 途中結果を複数生成 → 評価 → 有望なものを複数選択
→ 途中結果を複数生成 → ... → 評価 → 最終回答を選択



AIエージェント

■ **AIエージェント**: 目的を与えると、環境から情報を得ながら、推論・行動・観測を繰り返して目的を達成するAIシステム

- 行動 a とその結果得られる観測 o が含まれる

通常の推論: $x \rightarrow z_1 \rightarrow z_2 \rightarrow \dots \rightarrow \hat{y}$

AIエージェント: $x \rightarrow z_1 \rightarrow a_1 \rightarrow o_1 \rightarrow z_2 \rightarrow a_2 \rightarrow o_3 \rightarrow \dots \rightarrow \hat{y}$

■ **例**: 東京出張のホテル探し

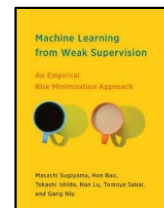
- x : 「来週の東京出張に最適なホテルを探して」
- z_1 : 条件を整理 (場所・予算・日程・評価)
- a_1 : Webでホテルを検索
- o_1 : 検索結果を取得
- z_2 : 価格・場所・評価を比較
- a_2 : 候補ホテルの詳細を確認
- o_2 : 空室・料金を取得
- $\hat{y}$: 最適なホテルを提案

手前味噌な研究：事後学習の改善 9

■ 教師付きデータの収集が困難：

- 弱教師付き学習

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



■ 教師情報に雑音が含まれる：

- ラベル雑音ロバスト学習

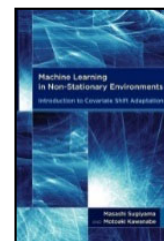
Niu, Han, Liu, Lu & Sugiyama.
Machine Learning with Noisy Labels,
MIT Press, in preparation



■ 環境が時間とともに変化する：

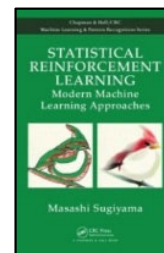
- 分布シフト適応学習

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

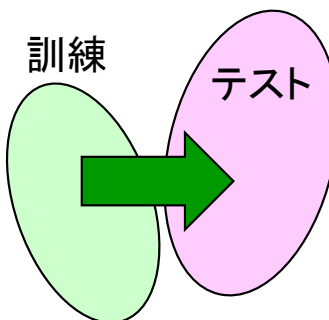
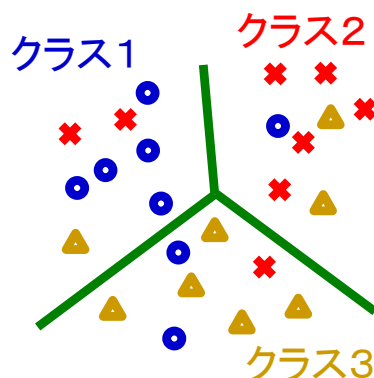
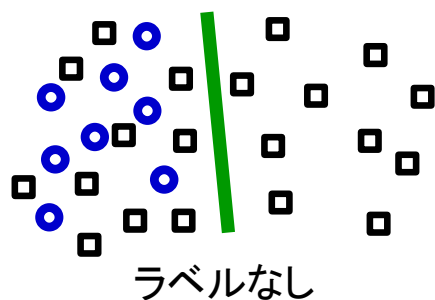


■ エージェントAIにも適用可能

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



正

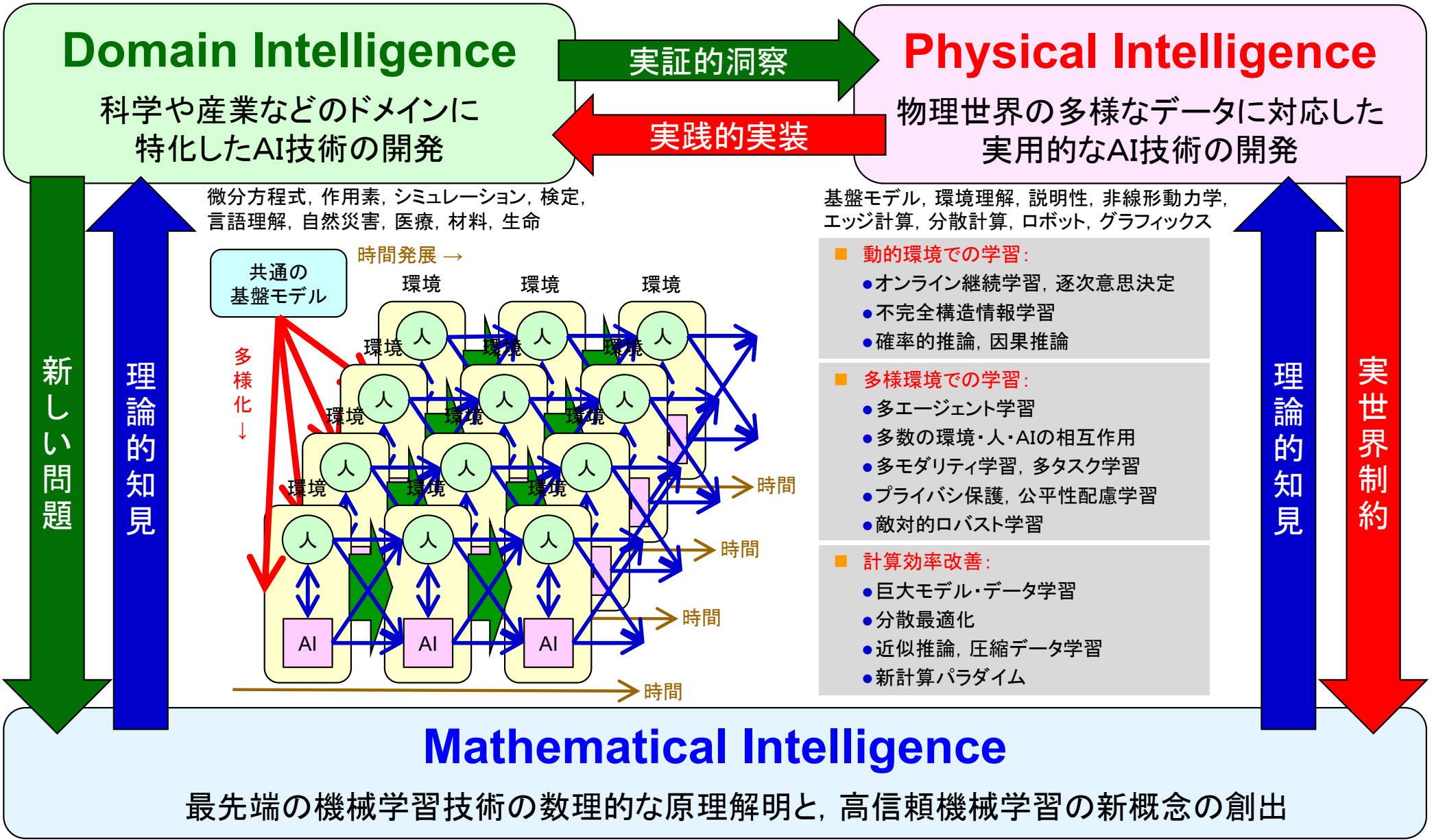


まとめ

- 古典的教師付き学習: $\hat{y} = f(x; \theta)$
 - タスクごとに専用モデルを学習
- 基盤モデル(事前学習): $\hat{y} = F(x, c; \hat{\Theta})$
 - 汎用モデルをタスクに適用
- 基盤モデル(事後学習): $\hat{\Theta} \rightarrow \hat{\Theta}'$
 - 人間の指示や好みに合うように追加学習
- 基盤モデル(テスト時計算): $\hat{y} = F(x, z_1, \dots, z_D, c; \hat{\Theta}')$
 - 推論・探索・検証で回答を改善
- AIエージェント: $x \rightarrow z_1 \rightarrow a_1 \rightarrow o_1 \rightarrow z_2 \rightarrow \dots \rightarrow \hat{y}$
 - 行動・観測しながら目標を達成
- 事後学習やAIエージェントは改善の余地がある:
 - 弱教師付き学習, ラベル雑音ロバスト学習, 分布シフト適応...

理研AIP:大規模基盤モデルを活用した動的多样環境での適応学習システムの開発

- 信頼性が保証される機械学習の数理的な理論体系を構築
- 物理法則に基づく高精度かつ低コストなシミュレーションを実現
- 実空間での省人化, 製造工程, 研究実験の全自動化に貢献



学習理論, 計算量, 近似計算, 情報理論, 動的システム, 統計力学, ゲーム理論