

2020年9月18日

資料 1

AI技術

中川裕志

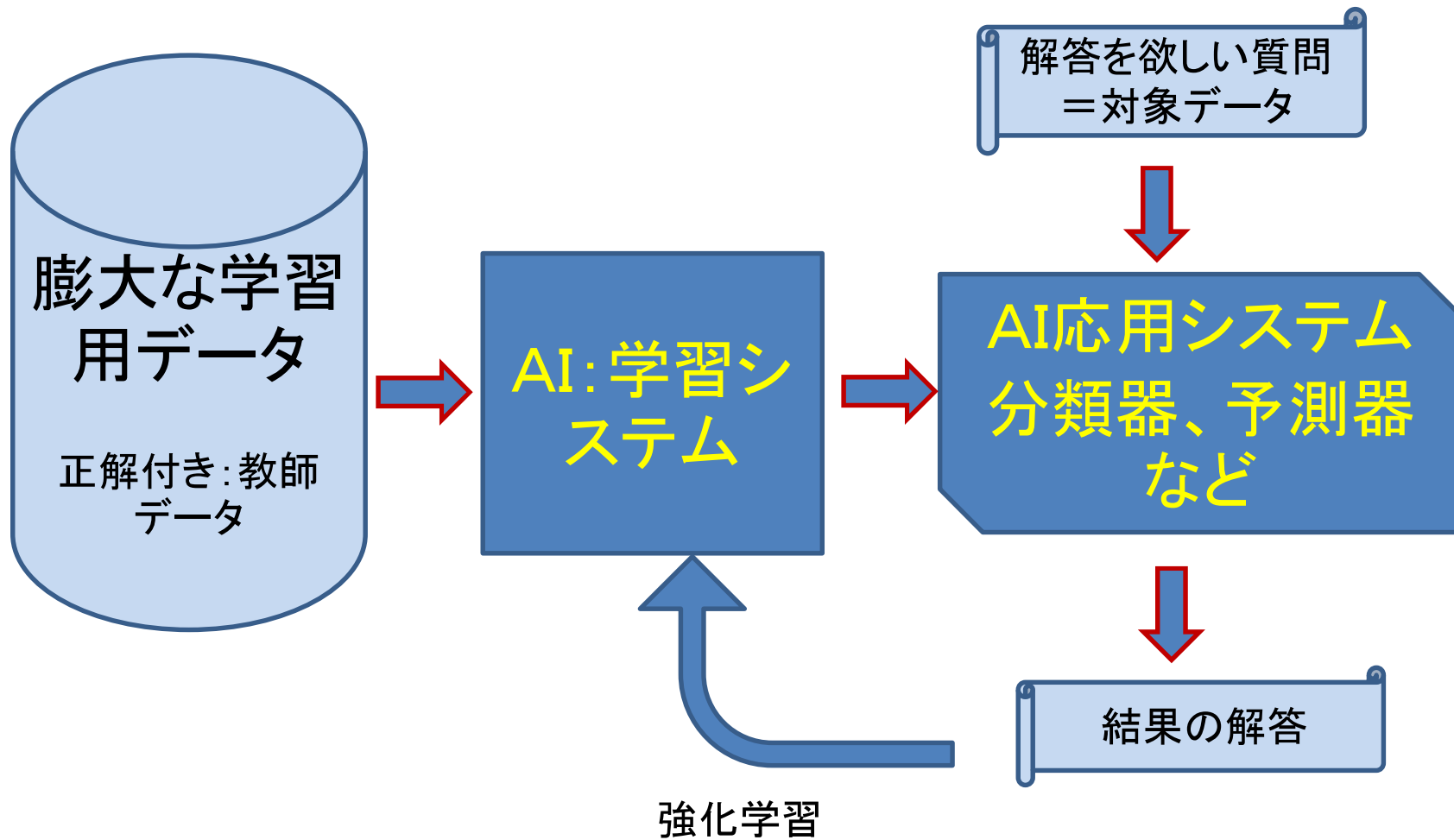
(理化学研究所・革新知能統合研究センター)

AI技術の用語

- 教師あり学習＝ラベルづけられた正解データから学習
- 教師なし学習＝ラベルなしデータを内容の類似性で分類
- 半教師あり学習＝教師あり学習の結果を使ってラベルなしデータを教師データ化する
- 強化学習＝使いながらその結果を使って学習し直し

- 古典的な統計的学習
 - 分類木、SVM、ランダムフォレスト、
 - 動作内容、性能予測が理論的にできる
- 深層学習
 - ビッグデータ勝負？ Attention
 - 動作内容、性能予測が理論的にできない

AIの仕組み：学習と応用システムの関係



人間はループの内か外か

- タスク実行において人間がループの内側にいるか外側にいるかという設計上の起点
- 内側
 - システムは人間の拡張。人間はシステムと共同作業
- 外側
 - システムは自律的に動き、人間はシステムに命令するだけ。データ収集さえ自動的。

Web広告の値付けー1

- どの場所にどの時刻に出すと効果的
- 効果～値段
 - ビッグデータの分析で精度が増す
 - 大規模プラットフォームほど機械学習で高い精度を出せる
- 個人対応 ターゲット広告
- 場所、時刻＋買ってもらえそうな商品
- 個人が買いたい商品を知る技術が難しい
- 協調フィルタリング

	商品1	商品2	商品3	商品4
顧客1	○	○		○
顧客2		○	○	
顧客3	○	○		広告出す

Web広告を表示する位置

- ブラウザ上でどの(ような)閲覧者に対して、どの日時にどの位置に広告を表示すればCTRやさらには買ってもらえるかというデータをプラットフォームは保持。
- したがって、広告場所の日時、場所で、ないし閲覧者に依存して広告料金が異なる。

検索エンジンの挙動調査

◆ 検索エンジンのランキングは

1. そのページにリンクしているページが多ければ、順位が上がる
2. 多くの検索がされたページは順位が上がる
3. 個人が過去に検索したページに類似したページがその個人に対してはランクが上がる

参考資料

- ◆ JST CRDSの「研究開発の俯瞰報告書」2019
https://www.jst.go.jp/crds/pdf/2018/FR/CRDS-FY2018-FR-02/CRDS-FY2018-FR-02_06.pdf
- ◆ EU AI白書
<https://home.jeita.or.jp/cgi-bin/topics/detail.cgi?n=2693&ca=13&ca2=76>
- Guidance for Regulation of Artificial Intelligence Applications: USA
Whitehouse. MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES (Draft 2019/4/24)
https://www.whitehouse.gov/wp-content/uploads/2020/01/Draft-OMB-Memo-on-Regulation-of-AI-1-7-19.pdf?utm_source=morning_brew
- ◆ QA4AI AIプロダクツ品質保証ガイドライン
<http://www.qa4ai.jp/QA4AI.Guideline.201905.pdf>
- ◆ AIと社会と法: 有斐閣(2020年8月)

ご清聴ありがとうございました

付録：Multi-stakeholder

Multi-stakeholder

- 人間がタスクの内側、外側のいずれの場合でも関係者は多い
 - 関係者： Multi-stakeholder
 - AI開発者
 - AIへ学習に使う素材データを提供した者
 - AIソフトを宣伝、販売した者
 - AIを販売、宣伝などに利用する者：プラットフォーム
 - 消費者
 - ITリテラシーがない消費者を100%保護すべきか？
 - どの程度のITリテラシーを消費者に期待すべきか？

付録：説明可能なAI

説明可能性: eXplainable AI: XAI

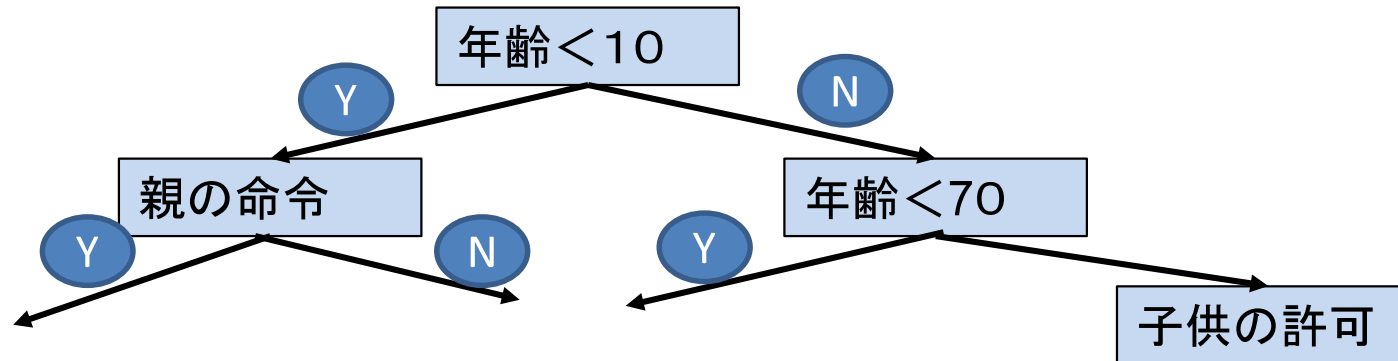
◆ブラックボックスAIの内部動作を多数の変数の値の変化で直接説明しても理解不可能

➤理解可能なAIで近似

– 決定リスト

- if 年齢 < 10 then 売らない
- else if 年齢 < 20 then 親の命令なら売る
- else if 年齢 < 70 then 売る
- else if 年齢 < 90 then 子供の許可があれば売る
- else 売らない

説明可能性: eXplainable AI: XAI

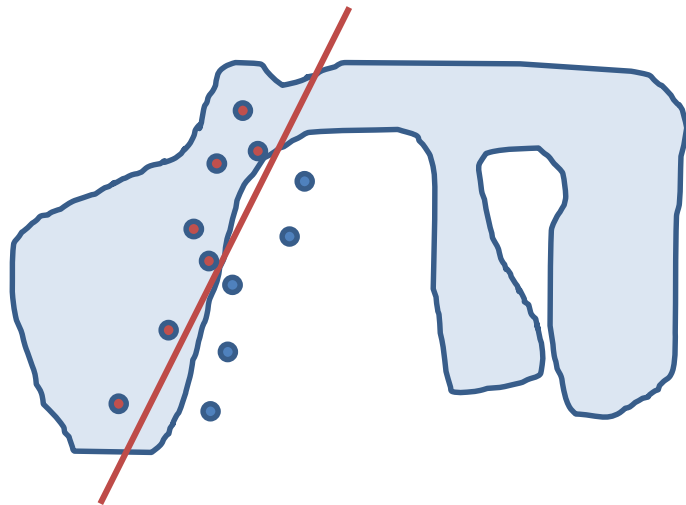


- ◆ 分かりやすい分類木で近似
- ◆ 類似例の結果を表示して理解してもらう
 - ・ 類似性の定義
 - ・ 恣意的な類似例でないことの保証
- 今までにカバーできていない例を内挿、外挿できることの説明
- ただし、あくまで粗い近似なので、分類木自体は証拠性に欠ける。
- 頼れるのは実例データのみ。

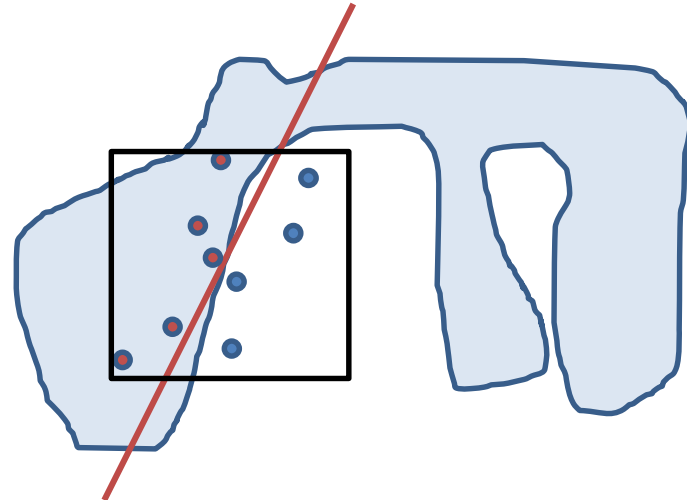
AI内部に動作説明能力を組み込む XAI(eXplainable AI)

- AIの動きを狭い領域に絞って
- $y = a_1 \cdot x_1 + a_2 \cdot x_2 + \dots$ で近似: LIME, SHAP, Anchor
- <https://www.slideshare.net/SatoshiHara3/ss-126157179?fbclid=IwAR1W8IREvEqIaKC1CZ74dEFDMqsWglsSjyuhO2LXyk8Gjt5j8U-OnJEn0s0>

LIME
局所的に境界線



Anchor
特徴量がどの範囲にあれば十分な
制度の境界線になるかが分かる



AI内部に動作説明能力を組み込む XAI(eXplainable AI)

- 結果に対する特徴量の寄与率を表示してくれるだけ
- ビッグデータで高次元特徴の場合、特徴量の意味を理解したうえで寄与率の値を客観的に評価できるか。