

AIの仕組み・歴史

- 人工知能(Artificial Intelligence)
- 研究者によって定義が大きく異なる。

人間の知的行為を計算機で模倣させたソフトウェア

- 現時点でも新たなプロダクト創出の可能性
- AIは本当に必要？

日本には必要 人が足りない、ノウハウが消失、効率性に課題 の為

- AIの歴史
 - 第一次AIブーム

1960-1970年代 推論と探索 ゲームへのトライアル

- 第二次AIブーム

1980-1990年代 知識工学、エキスパートシステム 専門家の代わりを人工知能にやらせる

- 第三次AIブーム

2010年代から 条件が揃ってきた3回目のブーム ハードウェアx最新技術xビッグデータ
DeepLearning

- AIの種類/仕組み

- 特化型人工知能

職人志向 1つの人間の知的行為のみに特化して動作 特定の処理においてのみ高い予測精度を発揮 昨今のAIの99%

- データ

AIは大量のデータを材料に規則性やルールを学ぶ その為、学習対象となるデータの量と質が重要

- 汎用型人工知能

人間と同じように新たな知識を自ら獲得 人間と同じような知的処理を発揮する能力を保有

- 知識表現 * イライザ(ELIZA)

1964~1966 ジョセフ・ワイゼンバウムによって開発。本物の人間と対話しているような錯覚(イライザ効果)に陥らせた。

- DENDRAL

1960年代にスタンフォード大学おエドワード・ファイゲンバウムによって開発 未知の有機化合物を特定するエキスパートシステム。世界初のエキスパートシステムと言われている。

- マイシン(MYCIN)

1970年代にスタンフォード大学で開発された。血液中のバクテリアの診断支援をする、ルールベースのプログラム。初期のエキスパートシステムとして最も影響力が大きかった。

- 意味ネットワーク - もともとは認知心理学における長期記憶の構造モデル。人工知能においても重要な知識表現の一つ

- 「is-a」の関係

継承関係。推移律が成立する。

- 「part-of」の関係

全体と部分の関係。推移律は必ずしも成立しない。最低5種類の関係がある。

AIと機械学習

- どうやってAIに知的学習をさせるのか

- 機械学習(Machine Learning)

データからパターンを学習し、パターンを発見したり結果を利用して将来を予測する手法のこと 確立や行列演算等、数学的アプローチを利用する

- 機械学習でできること

- 未知の情報を予測すること
 - 仲間分けをすること

- 機械学習と統計学

厳密な違いがあるとは言えないが、データ分析の志向性に違い

- 機械学習

未知への予測精度(汎化性能)を追及している傾向。高度なパターン認識="機械が立てた仮説"と考えられる。

- 統計学

データの性質の理解や説明を追及している傾向。特に検定等は"人が立てた仮説"を検証することが目的、データの性質の一般化を重視。

- 機械学習と深層学習

- 深層学習(Deep Learning)

Neural Networkと呼ばれる機械学習手法の発展形。昔は2~3層だったが、今は多いもので1000層を超えるハードウェア技術の進化と結びつき実現 特徴量加工 ⇒ 構造設計 特に画像/動画データに対して既存手法よりも高い精度を誇る

AIでは何ができるのか？

- ビジネスでの活用例

- 回帰(数量予測)

- 商品の需要予測
 - 目的：食品の需要予測をし、売り上げを最適化
 - 予測：次の日の食品の売り上げ数 * 着雪量予測 * 目的：列車への着雪量を予測し、人員配置最適化 * 予測：次の日の各時間帯の列車への着雪量
- 分類(ラベル予測)
 - ネット広告のクリック予測
 - 目的：ネット広告の効果最大化・最適化
 - 予測：ユーザx広告においてクリックするか否か
 - カード不正使用検知
 - 目的：クレジットカードの詐欺利用の防止
 - 予測：詐欺の疑いがあるか否かをログデータから予測
- テキスト
 - 自動校閲
 - 目的：校閲者の作業工数・負担の軽減
 - 予測：NGワードや誤字脱字を自動検出・指摘
 - メディカルコーディング
 - 目的：用語統一作業の自動化による作業効率化
 - 予測：入力文章に対して置換用語候補の出力
- 画像・動画
 - 画像・動画の物体検出
 - 目的：作業者の作業負担軽減及び品質向上
 - 予測：画像中の異物の検知
 - リアルタイム道路認識
 - 目的：自動運転の為にリアルタイム道路物体認識
 - 予測：道路上のオブジェクト名と領域を自動出力

AIクラウドサービス

- AI活用：3つの必要条件
 1. 計算資源
 2. データサイエンティスト
 3. データ
- 計算資源
 - AI学習の効率化のためには、HW、電源管理、温度管理等、お金と手間がかかる
 - AI特化型クラウド・コンピューティング・サービス
 - AI機能を簡単に利用できるAPI機能提供
 - データ前処理～モデル学習・評価をシームレスに実装可能な基盤を提供
 - 計算資源も併せて提供

1. Servicesレイヤー (SaaS)

物体認識、言語処理、音声等あらかじめ用意されたAIをAPIで利用

2. Platformレイヤー (PaaS)

AI開発からデプロイまでの機能

3. Infrastructureレイヤー (IaaS)

CPU/GPU、Storage等

データサイエンティスト

- データサイエンティストとは
 - データ分析を武器に現状分析からAIアルゴリズム開発までを担う専門家

1. ビジネス力

課題背景を理解したうえでビジネス課題を整理し解決する力

2. データサイエンス力

情報処理、人工知能、統計学などの情報科学系の知恵を理解し使う力

3. データエンジニアリング力

データサイエンスを意味の合う形に使えるようにし、実践、運用できるようにする力

- データサイエンティストの道具
 - 分析ツール・プログラム言語
 - Python、R、Jupyter Notebook等
 - 機械学習フレームワーク
 - 機械学習プログラムを実装するうえで便利なコードがまとめられたもの
 - 通常、分析者はフレームワークを利用してデータ分析及び実装を行う
 - ほとんどがOSSであり、利用目的を問わず自由に利用できる。
 - DeepLearning用フレームワーク
 - TensorFlow、PYTORCH、Chainer、Keras等

データ

- AIに必要なデータ
 - 教師あり学習
 - 入力データに対する正解が存在する
 - 目的変数/教師ラベル：予測したいもの
 - 説明変数：予測のヒントになりそうなもの
 - 意図したAIを作るためには"質の良い"学習データが重要
 - データからどんなAIが作れるかと考えるのではなく、作りたいAIに必要なデータとは何か？考えるのが望ましい。
- データを用意する方法
 - 自社データを活用
 - 公開データ（有償・無償）を活用
 - データを独自に作る
- データを独自で作る場合の注意点
 - どんなAIを作り、どう利用していくかを念頭に考え、戦略的にデータおw取得していく