

AIエンジニアリング (1)

機械学習とは？

わたしについて

経歴

- 五木田 和也 (ごきた かずや)
- Twitter: @kazoo04
- 2006年頃～ 機械学習が趣味になる
- 2010年～ 機械学習が研究テーマになる
- 2012年～ 機械学習が仕事になる



Usageee

GARM
ガール合同会社

Repro

所属



Usageee

(株)ウサギィ

取締役

AI関連の研究開発

アプリケーション開発

GARM

ガルム合同会社

ガルム(同)

CEO

汎用AIの研究開発

ロボット制御

Repro

Repro(株)

AI顧問

マーケティング
プラットフォーム

その他の活動(汎用人工知能系)



連載：

脳から始める汎用AI(日経BP)



全脳知能アーキテクチャ・イニシアティブ
正会員

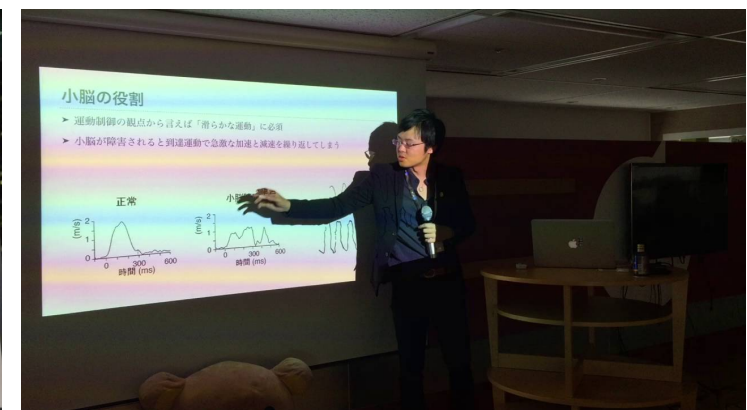


著書：

コンピュータで「脳」がつくれるか(技術評論社)



日本ディープラーニング協会
認定プログラム



この科目について

この科目について

- 目標:
 - 機械学習の知識をつけ、Pythonで実装できる
 - ほぼ画像認識、自然言語処理
- 評価方法:
 - 中間テスト(25%)
 - 中間レポート(25%)
 - 最終テスト(25%)
 - 最終レポート(25%)

そもそも人工知能(AI)とはなんだろう

AIの定義

- 実は統一的な見解・定義はほぼないと言ってよい
- 人工知能学会のサイトですら粒度がバラバラでうまく分類できていない

遺伝アルゴリズム
エキスパートシステム
音声認識
画像認識
感性処理
機械学習
ゲーム
情報検索
推論

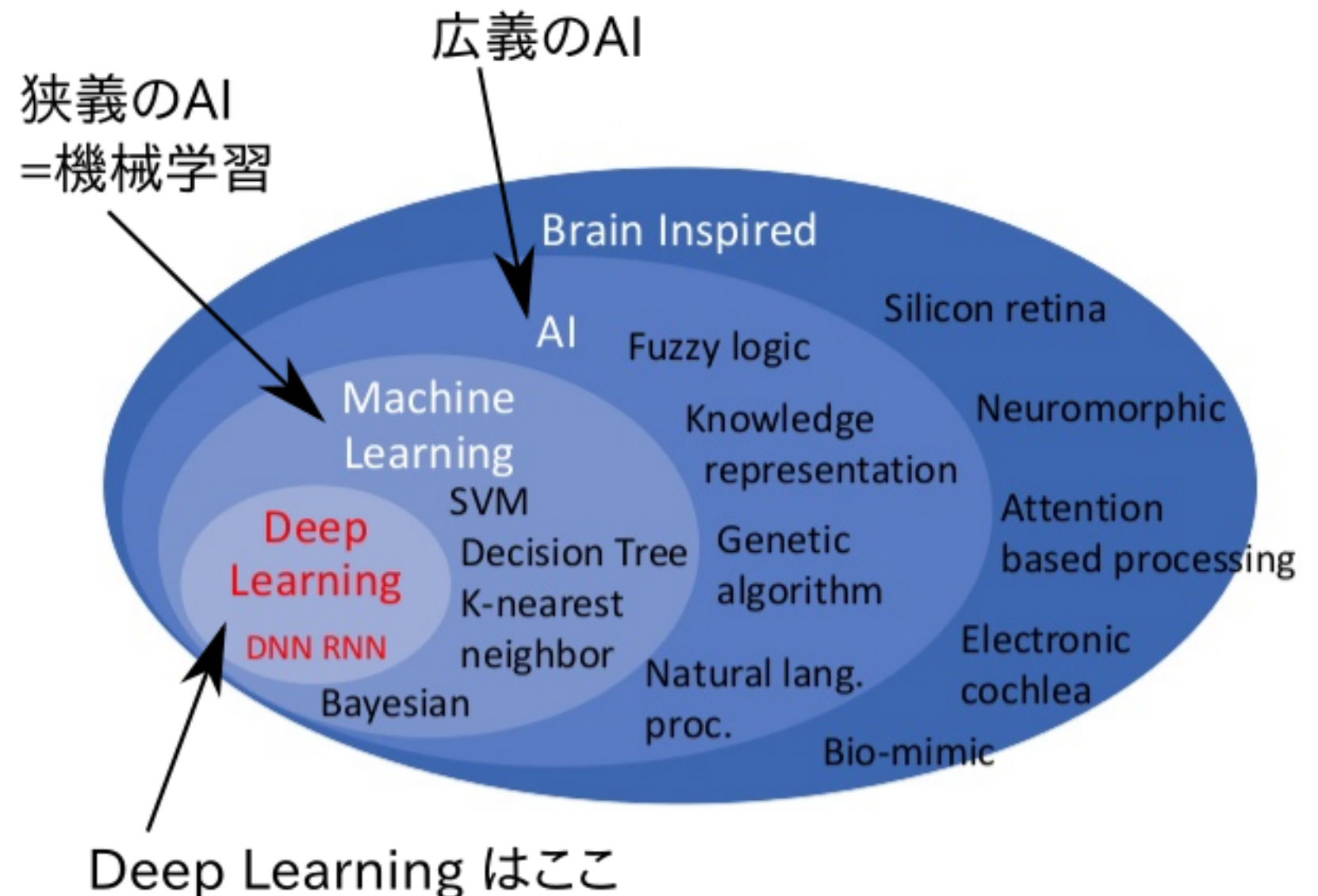
探索
知識表現
データマイニング
ニューラルネット
ヒューマンインタフェース
プランニング
マルチエージェント
ロボット工学

人工知能学会: What's AI? より

AIの定義

- 「知的な動作ができるプログラム」はトートロジー
- その時代において知的な振る舞いをしていると思えるのならAI
- 最近は...

機械学習 = AI



この業界で大事なこと

- 機械学習のエンジニア・研究者に「人工知能(AI)」と
言うてはいけない
- この科目名も「AIエンジニアリング」だが、実際には機械学習 (Machine Learning) を扱う科目なので...
- 今後は「人工知能」、「AI」の用語は特別な場合を除いて使わない

機械学習とはなんだろう

機械学習とは

- 実は機械学習の定義もあまりはっきりしたものはない
- ただ、一般的に機械学習は次の3種類からなるので、この総称とっていい
 - 1. 教師あり学習
 - 2. 教師なし学習
 - 3. 強化学習

1.教師あり学習(supervised learning)

1. 人間がお手本(教師データ)を用意してあげる
 2. 教師データをもとにして、人間と同じように判断できるまで学習する
- 例
 - 迷惑メール(スパムメール)フィルタ
 - 文字認識
 - 画像認識
 - 音声認識

2.教師なし学習(unsupervised learning)

1. 人間がデータを用意する
 2. データからなにかしらの法則などを見つけ出す
- 例
 - データマイニング系
 - クラスタリングなど
 - 機械学習の中では下火だったが、最近アツい分野

3. 強化学習(reinforcement learning)

1. 直接お手本は用意しないが、人間がゴールだけ決める
 2. あとは試行錯誤して勝手にうまくなっていく
- 例
 - ゲームAI
 - 囲碁、将棋、チェスなど
 - ロボット制御
 - 一見便利そうに思えるが、学習がすごく難しいので教師あり学習のほうが便利ことが多い

教師あり学習を体験してみよう

その前に...

- 教師あり学習がやっていることを手動でやってみよう

Boston Dataset

- ボストンの住宅価格を予想する問題
- 犯罪発生数や部屋の平均数など14の項目から予想する

項目名	説明
CRIM	人口 1 人当たりの犯罪発生数
ZN	25,000 平方フィート以上の住居区画の占める割
INDUS	小売業以外の商業が占める面積の割合
CHAS	1:チャールズ川が近い 0:川が近くない
NOX	NOx の濃度
RM	住居の平均部屋数
AGE	1940 年より前に建てられた物件の割合
DIS	5つのボストン市の雇用施設からの距離
RAD	環状高速道路へのアクセスしやすさ
TAX	\$10,000 ドルあたりの不動産税率の総計
PTRATIO	町毎の児童と教師の比率
B	町毎の黒人 (Bk) の比率を次の式で表したもの。
LSTAT	給与の低い職業に従事する人口の割合 (%)



このへん



どうやって予想するか？

- とりあえず次のような式を考えて、そこから予想しよう
- $A * CRIM + B * ZN + \dots + M * LSTAT + N = \text{住宅価格}$
- みたいなのを考えて、 $A \sim N$ までの値をいい感じに調整すれば、住宅価格を予想する式ができそう
- こういうのをモデルという
 - たとえば「今回は住宅価格を予測するためにこのようなモデルを考えます」のように言う

どうやって予想するか？

- 今はいりあえず各属性にそれぞれ1個ずつ数字を掛け算するようなモデルを考えた
- 特に正解はないので好きなモデルを考えれば良い
- ただ、この形式のモデルは一番単純なのに一番奥が深いのでよく使われる
- こいつはよく使うので特別に名前がついていて、線形モデル(線形回帰モデル)という