

人工知能技術について

東京大学大学院 情報理工学系研究科

原田牛久研究室 修士2年

福田圭佑

自己紹介

- 東京大学大学院 情報理工学系研究科 修士2年
- 原田牛久研究室 所属
 - 専門は画像認識
- 先端人工知能寄付講座 最優秀賞頂きました！
 - 引き分けを目指すオセロAIの作成



Agenda

1. 人工知能とは？
2. 人工知能技術の歴史
3. 近年の人工知能技術 (主に機械学習) の基礎
4. 産業へのインパクト

人工知能とは？

究極的には、

「人間と同等に知的な振る舞いを見せるコンピュータ」

人間の脳は電気信号 + 化学反応なんだから
プログラムでかならず実現可能という発想



だが、そんなものは当然まだできていない!!!

人工知能とは？

- 専門家たちの意見も様々
 - 知能をどう捉えるかによってばらつき
- 個人的にしつくりきているのは
「その時代において最も傑出した情報技術」
 - OCR、かな漢字変換もかつては人工知能だったはず
 - コモディティ化し当たり前のものとなると、そう呼ばれなくなる
 - 自分の理解が及ばない情報技術は大体人工知能に見える

図1 専門家による人工知能の定義

中島秀之 公立はこだて未来大学学長	人工的につくられた、知能を持つ実体。あるいはそれをつくることによって知能自体を研究する分野である
西田豊明 京都大学大学院 情報学研究科教授	「知能を持つメカ」ないしは「心を持つメカ」である
溝口理一郎 北陸先端科学技術 大学院大学教授	人工的につくった知的な振る舞いをするもの（システム）である
長尾 真 京都大学名誉教授 前国立国会図書館長	人間の頭脳活動を極限までシミュレートするシステムである
堀 浩一 東京大学大学院 工学系研究科教授	人工的につくる新しい知能の世界である
浅田 稔 大阪大学大学院 工学研究科教授	知能の定義が明確でないので、人工知能を明確に定義できない
松原 仁 公立はこだて未来大学教授	究極には人間と区別がつかない人工的な知能のこと
武田英明 国立情報学研究所教授	人工的につくられた、知能を持つ実体。あるいはそれをつくることによって知能自体を研究する分野である（中島氏と同じ）
池上高志 東京大学大学院 総合文化研究科教授	自然にわれわれがペットや人に接触するような、情動と冗談に満ちた相互作用を、物理法則に関係なく、あるいは逆らって、人工的につくり出せるシステムを、人工知能と定義する。分析的にわかりたいのではなく、会話したり付き合うことで談話的にわかりたいと思うようなシステム。それが人工知能だ
山口高平 慶應義塾大学理工学部 教授	人の知的な振る舞いを模倣・支援・超越するための構成的システム
栗原 聰 電気通信大学大学院情報 システム学研究科教授	工学的につくられる知能であるが、その知能のレベルは人を超えているものを想像している
山川 宏 ドワンゴ人工知能研究所 所長	計算機知能のうちで、人間が直接・間接に設計する場合を人工知能と呼んでよいのではないかと思う
松尾 豊 東京大学大学院 工学系研究科准教授	人工的につくられた人間のような知能、ないしはそれをつくる技術

出典：『人工知能学会誌』より

人工知能とは？ - 分類 -

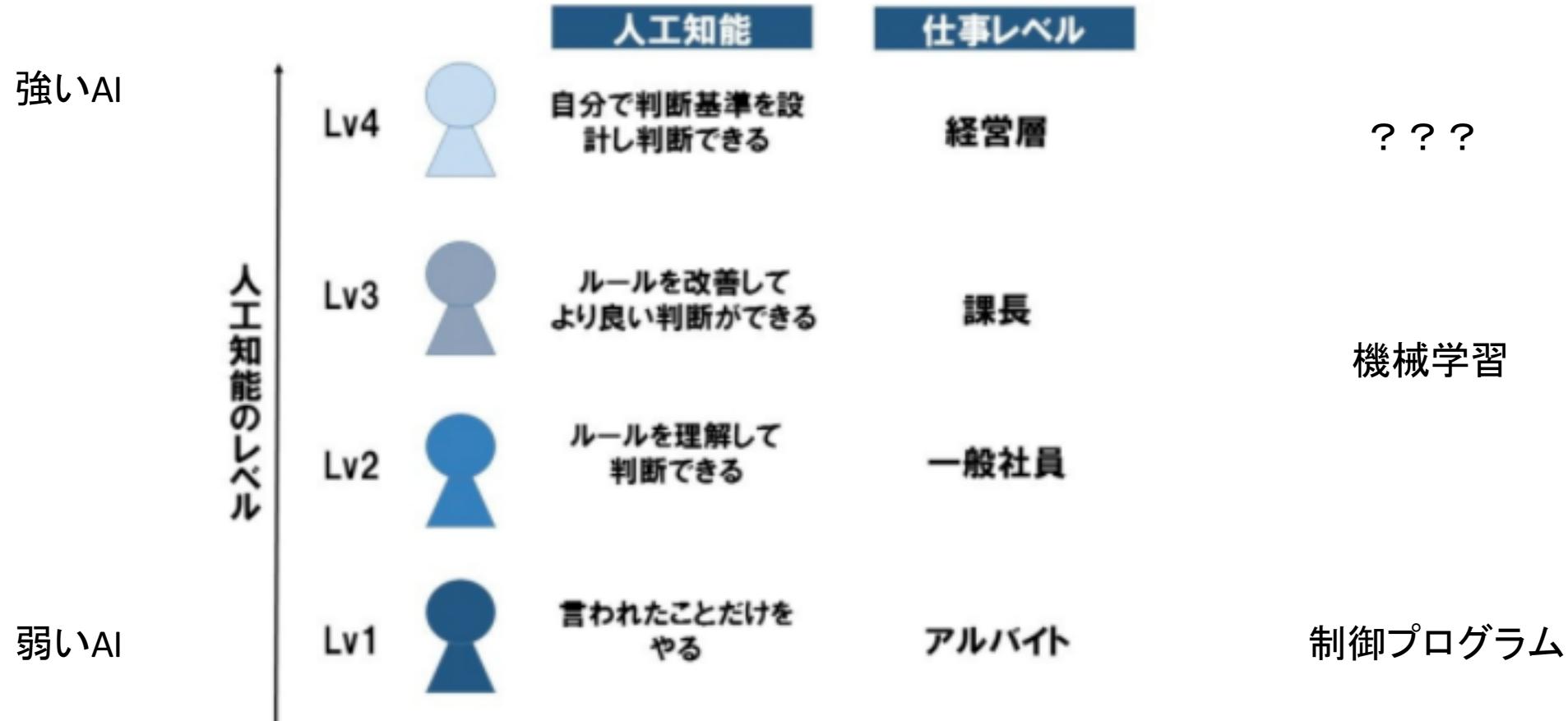
汎用人工知能

- 異なる領域で多様な複雑な問題を解決する知能
 - 設計時の想定を超えた
新たな問題に対しても対処可能
 - 自己理解・自律的自己制御
- AI当初の目的で、思い浮かべるのはこっちだが、現実から**あまりに遠いため**取り組みは少ない

特化型人工知能

- 特定の領域において知的に情報を処理する知能
- 既に**人以上**の能力が実用化
コンピュータ将棋・囲碁・チェス
画像認識, 自動運転, 医療診断
スパムフィルタリング
- 世間で言われてるのは**大体こっち**

人工知能とは？ - レベル -



何故今盛り上がってるのか？

- 大量のデータを得ることができる環境
 - 半導体技術の向上やデータベースの検索システムの速度向上
- 計算機能能力の向上
 - 一秒間に処理できる量の向上
- 様々な機械学習アルゴリズムの登場
 - Deep Learning等
 - パターン(文字、画像、音声等)認識精度の向上

3つが合わさって、ようやく実用的な技術になってきたということ

Agenda

1. 人工知能とは？
2. 人工知能技術の歴史
3. 近年の人工知能技術 (主に機械学習) の基礎
4. 産業へのインパクト

人工知能の歴史

- 第1次AIブーム: 探索・推論の時代 (1956 ~ 1960年代)

- 人工知能 (Artificial Intelligence)という言葉の誕生
 - 数学の定理証明・チェスを指す人工知能

→ 冬の時代

考えるのが早い人工知能

- 第2次AIブーム: 知識処理の時代 (1980年代)

- エキスパートシステム
 - 医療診断、有機化合物の特定、etc..
 - 第5世代コンピュータプロジェクト

→ 冬の時代

ものしりな人工知能

- 第3次AIブーム : 機械学習の時代 (2000年代~)

- ウェブとビッグデータ
 - 計算機能力の向上
 - 機械学習アルゴリズムの発展

データから学習する人工知能

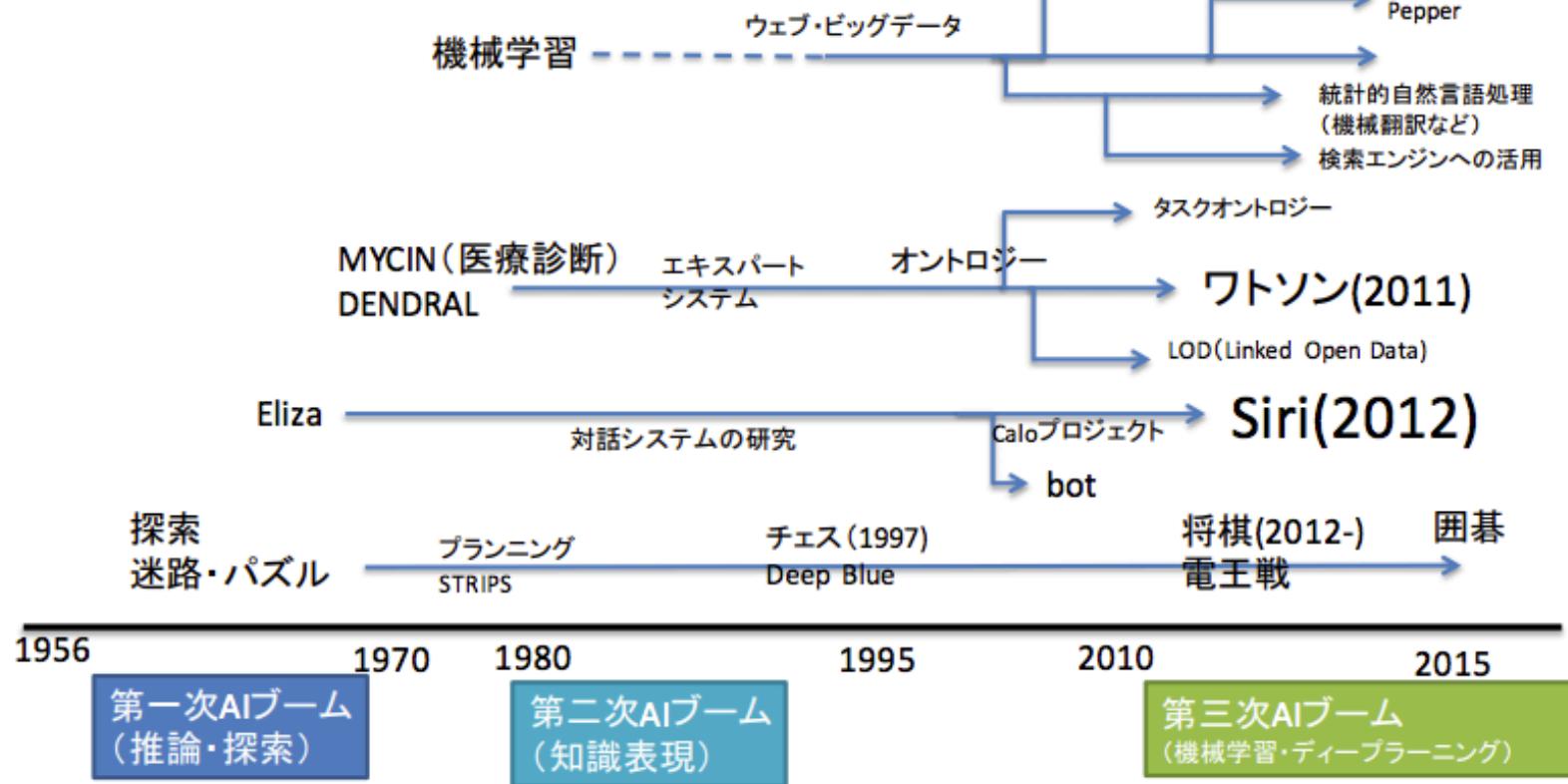


将棋電王戦



IBM ワトソン

<http://venturebeat.com/2011/02/15/bm-watson-jeopardy-2/>, <http://weekly.ascii.jp/elem/000/000/207410/>

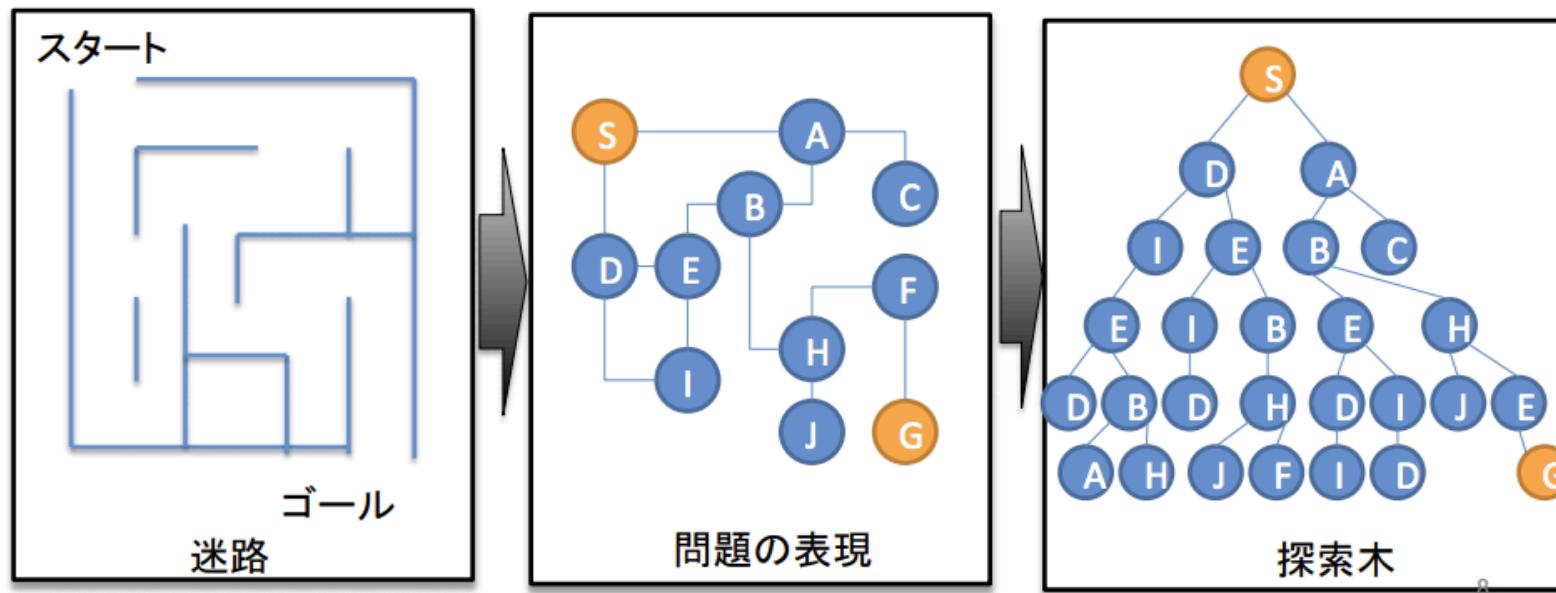


第1次AIブーム:探索・推論

- ダートマスワークショップ(1956)
 - 人工知能(Artificial Intelligence)という言葉が決まる
 - 世界初のコンピュータ(ENIAC)が出来た1946年からわずか10年後
- 第一次AIブーム
 - ゲーム・迷路を解くAI
 - 定理証明器(1957)
 - ニューラルネットワーク(1963)
 - 遺伝的アルゴリズム (1958)
 - DENDRALプログラム(1969)
 - 質量分光計の情報から分子構造を同定する
- 1970年～冬の時代

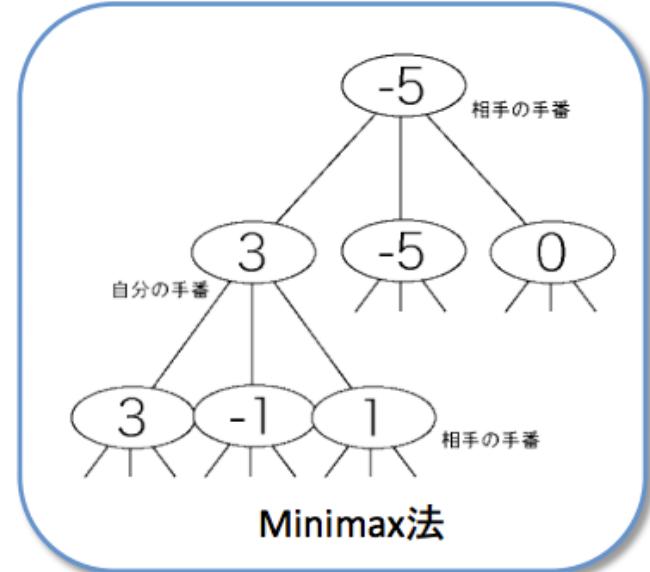
探索

- 探索木
 - Sからはじめて展開していき、Gが見つかれば探索成功
 - いろいろな戦略
 - 深さ優先探索: どんどん進む
 - 幅優先探索: 一段ずつ進む



ゲーム

- 盤面を評価する関数を設計する
- 相手も最善手を選択すると仮定して**探索** (Minimax法)
- 囲碁、将棋などの探索範囲が広すぎるゲームでは勝てず
 - しかし近年では、どのゲームでも人間に勝利
 - 計算速度の向上
 - 手法の進化 (モンテカルロ法など)
 - 盤面の評価関数の進化



第1次ブーム: 探索・推論 まとめ

- 解きたい問題を、探索・推論問題として、うまく記述すれば解ける
 - けど、逆に言えばきれいに記述できないと解けない
 - 結局トイプロblemしか解けず現実的な問題は解けない
-
- さらに、、、
 - 機械翻訳に対する否定的なレポート
 - (当時はメモリも速度も全然まだまで、20単語がメモリに乗るのが限界 (無理に決まってる)
 - MinskyによるPerceptronの限界証明

→なんだ、全然無理じゃん

→「人間の脳って電気信号なんだからコンピュータで実現できるっしょ！」という楽観的な期待が一気に失望に

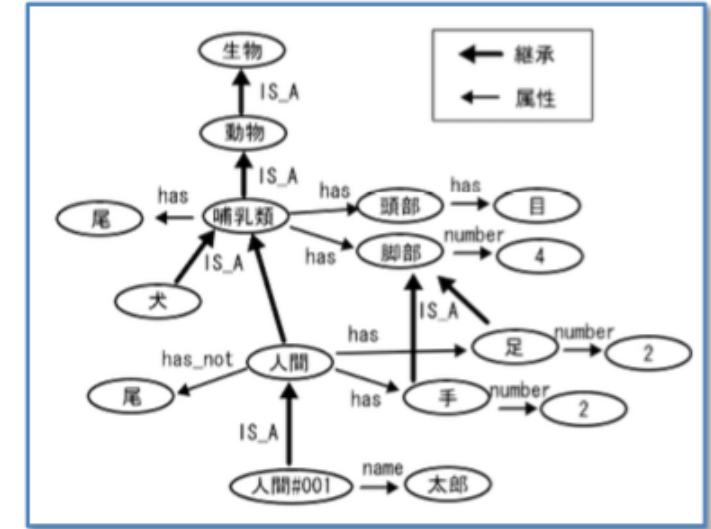
→1970年～冬の時代

第2次AIブーム:知識を入れると賢くなる

- 1980年～ 第二次AIブーム(知識の時代)
 - エキスパートシステム
 - 第5世代コンピュータプロジェクト(1981):570億円 国家プロジェクト
 - AIが産業へ
- 1995年～再びAIの冬の時代

知識表現

- 人工知能による人間の知識の表現手段
- Ex. 意味ネットワーク (1960-)
 - 人間の記憶の一種である意味記憶の構造を表すためのモデルの一種
 - ネットワーク型・フレーム型・ルール型などがある
- エキスパートシステム
 - 知識表現によって専門家の知識を記述して専門家の作業を代替させよう！
- 一般常識をデータベース化して人間と同等の推論システムを構築することを目的とするプロジェクトがあるが、20年たっても終わらない。。



意味ネットワーク

※ has関係は、part-of関係の逆

Watson

- IBMが開発した質問応答システム
- 2011年1月に米国クイズ番組 “Jeopardy!”で人間に勝利
- 現在は医療診断に応用されている

質問文:「本州のなかで最も西に位置するこの県は、1871年に発足した。」
正答:「山口(県)」



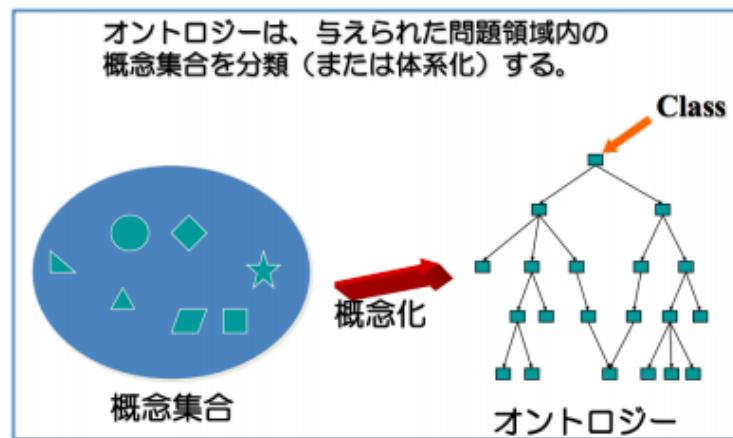
観点＼解候補	広島	山口	鳥取県	中国地方	奥多摩
候補と質問で型が一致する? (「県」である)	○	○	○	×	×
条件の一部が一致? (最も西にある)	×	○	×	○	○
時間表現が共通? (1871年の記述を含む)	×	○	×	○	×
該当する語句へのリンクの数 (多いほうがよい)	1300	500	200	150	10
総合点(確信度)	2%	92%	20%	6%	0%

表 -2 解候補ごとの根拠の探索

質問応答システムWatsonが示す未来、ProVISION, 2011

オントロジー

- エキスパートシステムのルールが数千、数万になると
 - 書ききれない
 - メンテナンスできない
 - 共有や再利用ができない→ 体系化しよう
- 概念化の明示的な仕様



ELIZA

- 対話システム
 - タスク志向型、非タスク志向型とあるがこれは後者、Siriは前者にある
- 対話にあたって ルールを記述し、患者と対話をを行う
 - “My head hurts” → "Why do you say your head hurts?"
 - “My mother hates me” → "Who else hates you?"
- ただ人手で与えられたルールに従って返事を出力してるだけなので最近では人工知能と対比して人工無脳とも呼ばれる
 - Twitter ボット

The screenshot shows a window titled 'File Edit Options Buffers Tools Help' with various icons above it. The main area contains a conversation transcript:

```
I am the psychotherapist. Please, describe your problems.  
Each time  
you are finished talking, type RET twice.  
  
I have a terrible headache.  
Maybe you should consult a doctor of medicine, I am a psychi-  
atrist.  
Okay, I'll do that then.  
Can you elaborate on that?  
What are you talking about? No I cannot elaborate. See you later.  
Why do you say that?  
Now you're just talking nonsense!  
What makes you believe now I am just talking nonsense?  
You are not making any sense.  
Perhaps your plans have something to do with this.  
You: ** *doctor* (Doctor Fill)--L1--Top-----15-----
```

フレーム問題

- 人工知能は現実に起こりうる全ての問題に対処することができない
- 「マクドナルドでハンバーガーを買え」と言わされたら
 - 現実世界では無数の出来事が起きる可能性があるが、ほとんどは今回のタスクには関係ない
 - タスクに関係あることだけを抽出しないといけない。
 - つまりフレームを作ってその枠だけで思考する
 - しかし、ある事象がこのタスクと関係あるかどうかを評価しようとしても、組合せ爆発によって無限の時間がかかる

シンボルグラウンディング問題

- 記号設置問題とも
- 記号システムにおけるシンボルが実世界の意味と結び付けられるかという問題
- 与えられた記号の意味を理解していないので、記号の操作だけでは知能を実現できないシンボルをそれが意味するものと結び付けるといけないが、非常に困難
- 馬とシマシマの意味を知っていれば、**シマウマ = 馬 + シマ** の関係を教えられれば、シマウマを見ればそれとわかるが、機械では無理



第2次AIブーム:知識処理 まとめ

- ・沢山覚えられるようになったので、知識を書けば前よりも賢くなつた!!
- ・でも、そのために知識を書くのがとても大変
- ・というか全てを書ききるのは無理では。。？
- ・+ フレーム問題・シンボルグラウンディング問題

→ 実用的なAIってやっぱり無理じゃない ???

→ 1995年～再びAIの冬の時代

第3次AIブーム:機械学習から表現学習へ

- 1990年~
 - 情報検索:米国政府主導によるTREC(Text Retrieval Conference)(1992)
 - データマイニング : 学会やジャーナルが増加
 - 検索エンジン:Google(1998)
- 2000年~
 - ウェブの広がり :1995年に 1万サイト、2006年に 1億サイト
 - ビッグデータ
 - 大量データを利用した機械学習の実用化
- 2012年~第三次AIブーム ←ココ

Agenda

1. 人工知能とは？
2. 人工知能技術の歴史
3. 近年の人工知能技術 (主に機械学習) の基礎
4. 産業へのインパクト

第3次AIブーム:機械学習から表現学習へ

- 大量のデータを得ることができる環境
- 計算機能力の向上
- 様々な機械学習アルゴリズムの登場

→これらにより、パターン(文字、画像、音声等)認識精度が**実用的なレベル**まで向上

→Google, Microsoft, Facebook, BaiduがこぞってAI研究所を設立

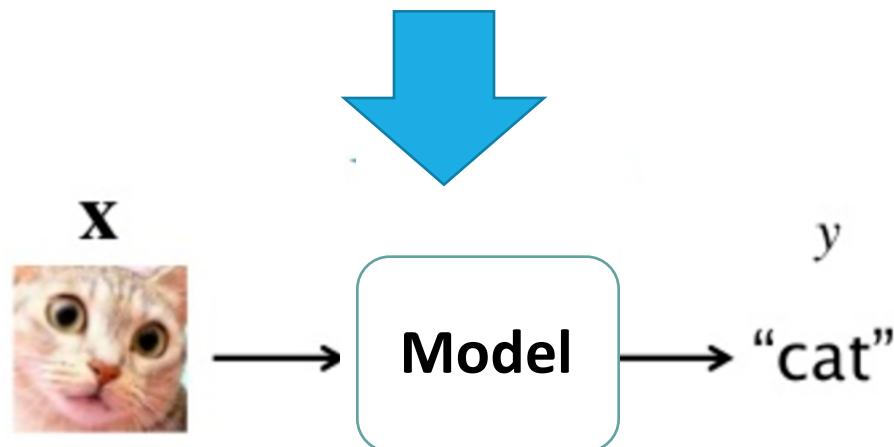
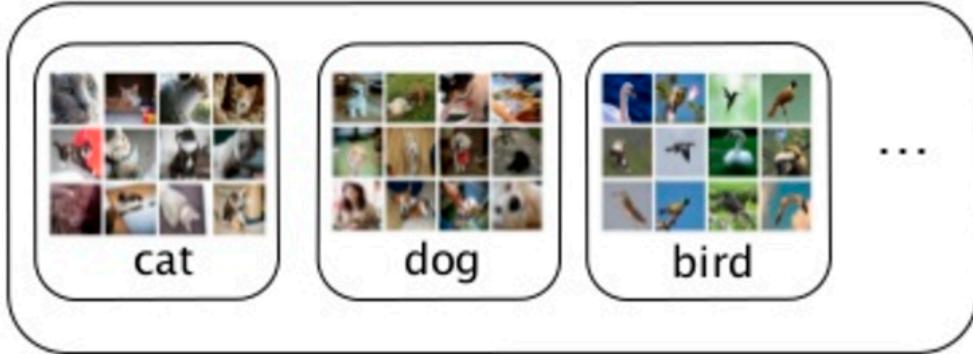
→大ブーム

機械学習 (Machine learning)

- 人工知能における研究課題の一つ
 - 人間が自然に行っている学習能力と同様の機能をコンピュータで実現しようとする技術・手法
-
- 教師あり学習
 - 入力とそれに対応すべき出力(ラベル)を出力する関数を学習する
 - 教師なし学習
 - 入力のみ (ラベルがついていないデータ)から何らかのモデルを構築する
 - クラスタリング、相関ルールの抽出

教師あり学習 (supervised learning)

- 入力と出力のペア (x_i, y_i) の対応を学習
($x \rightarrow y$ のマッピングとも言う)



未知データに対する予測を可能にすることが目標

機械学習の基本

1. 入力データから人手で特徴量を設計、抽出する

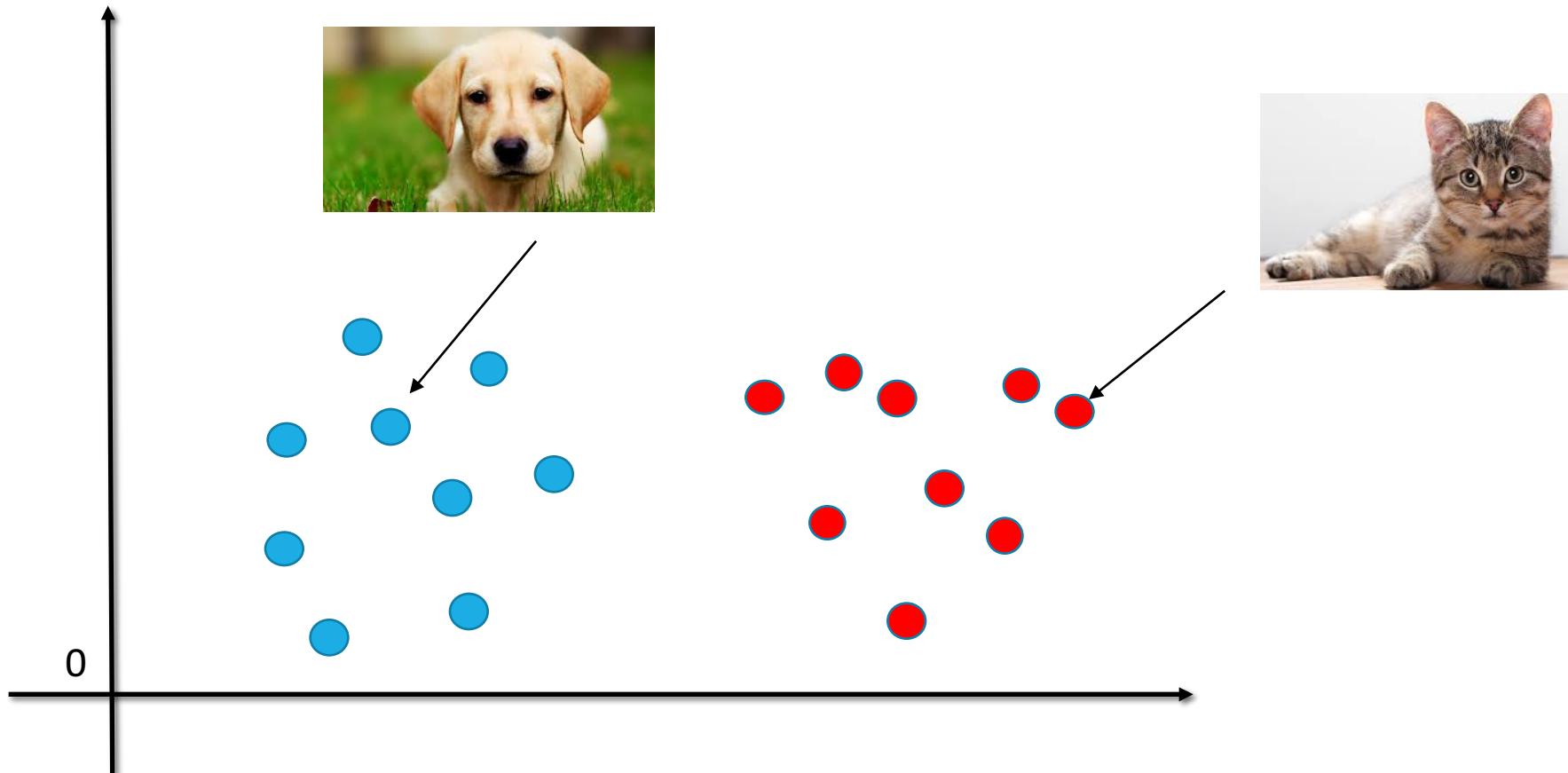
- 画像 : 色・テクスチャ・輝度勾配 etc..
- 音声 : 周波数解析
- テキスト : 似た意味の言葉を集計したり

2. 特徴量に対する学習・推論を行うモデルの学習

- 線形分離器 (SVM, Logistic回帰)
- 非線形分離器 (決定木, Kernel法, Bayes学習)
- サンプルベース (kNN)

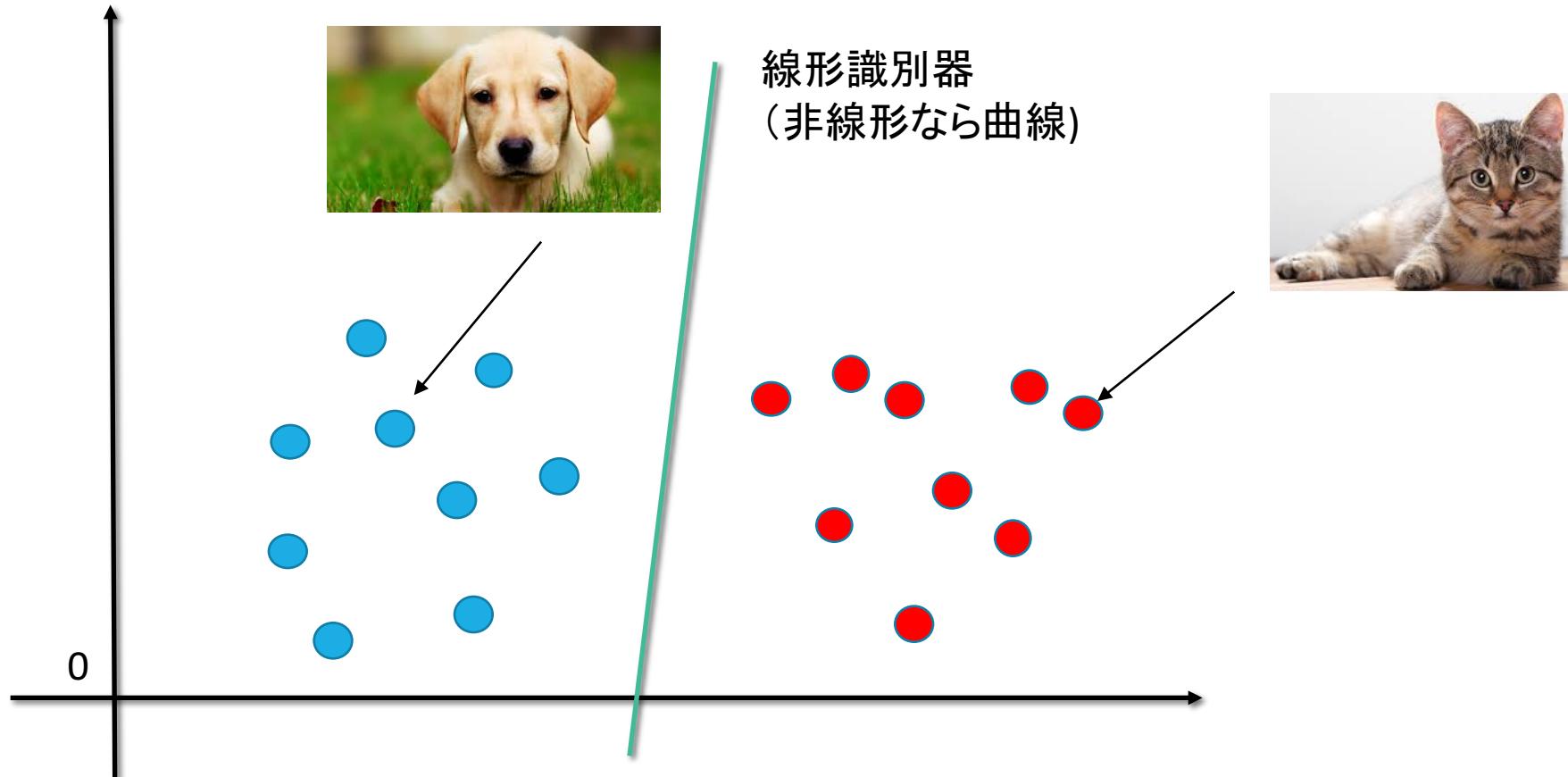
機械学習の基本

入力データから人手で特徴量を設計、抽出 (例えばここでは2次元の特徴量)



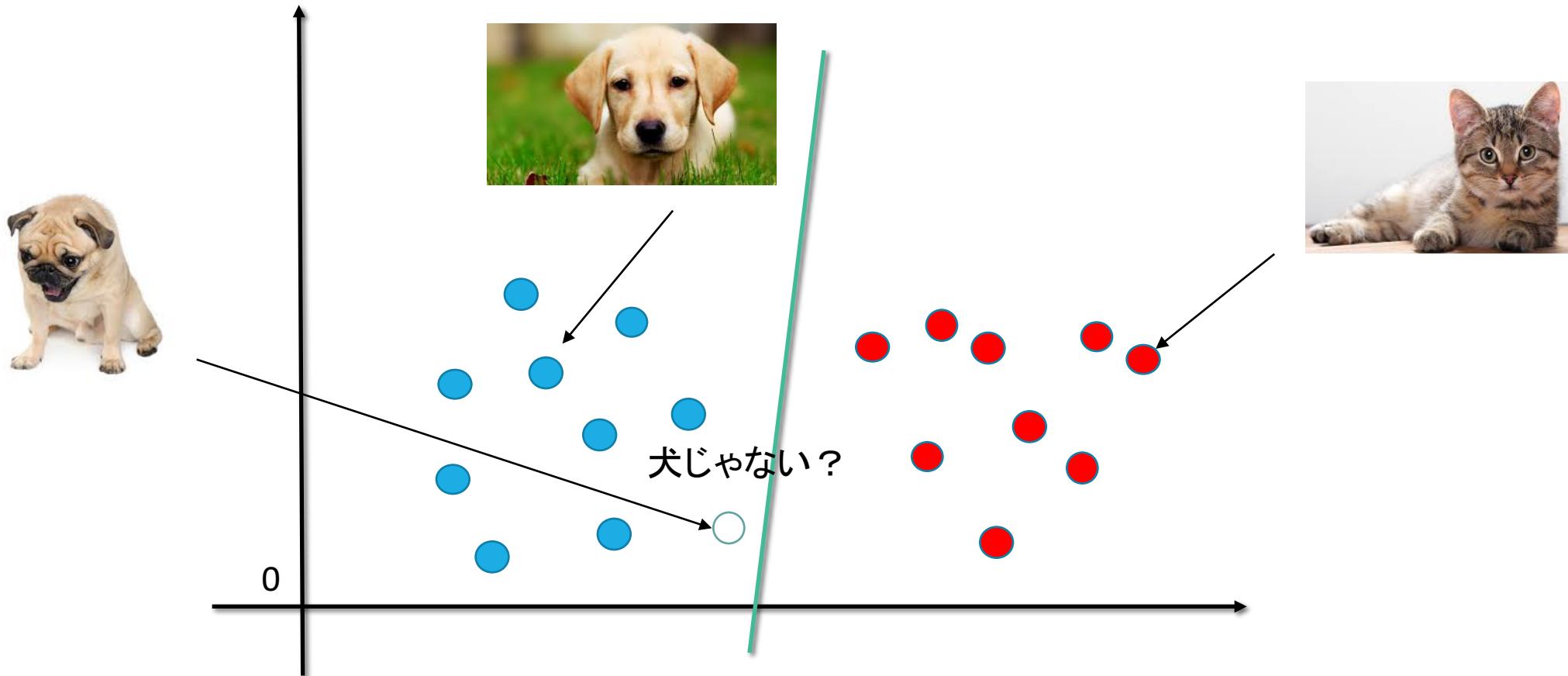
機械学習の基本

特徴量に対する学習・推論を行うモデルの学習（この場合、どう線を引くか）



機械学習の基本

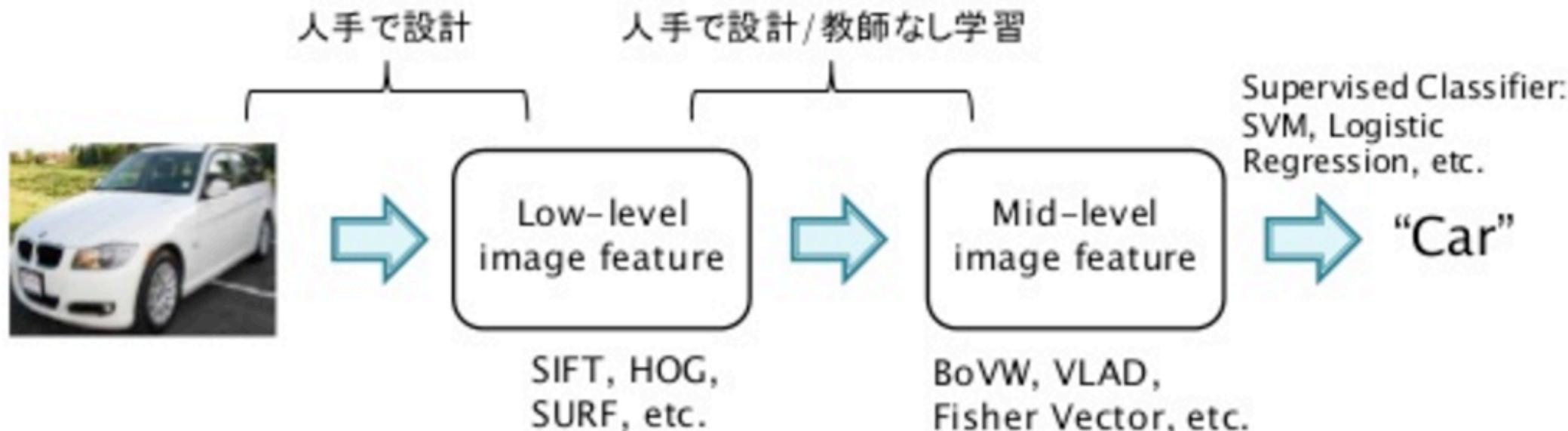
未知データに対する推論



例：一般画像認識

入力された画像に何が写っているかを当てるタスク

- ・ 線を引く時は曲線のほうが柔軟なんじゃないか
- ・ いや、曲線だと外れ値に弱いんじゃないか



- ・ 明るさの変化量で物体の形状がわかるんじゃないか
- ・ 二方向の直線があったらそれがコーナーなんじゃないか
- ・ 画像中にどういう特徴(エッジ、コーナー)がたくさんあるかをうまくまとめることができたら良い特徴になるんじゃないか

例：コンピュータ将棋

- 盤面の評価関数の学習
 - 盤面を入力して、その盤面がどれくらい良いかを学習
- 打ち手
 - 顔面を入力して、過去の棋譜での打ち手を学習
- 3駒の相対位置を利用するものが良いことが最近発見された
(1億次元くらい)

将棋の学習



人工知能 (機械学習)の課題

1. 入力データから人手で特徴量を設計、抽出する

- 画像 : 色・テクスチャ・輝度勾配 etc..
- 音声 : 周波数解析
- テキスト : 似た意味の言葉を集計したり

2. 特徴量に対する学習・推論を行うモデルの学習

- 線形分離器 (SVM, Logistic回帰)
- 非線形分離器 (決定木, Kernel法, Bayes学習)
- サンプルベース (kNN)

← こっちはある意味自動化できている

人工知能 (機械学習)の課題

1. 入力データから人手で特徴量を設計、抽出する

- 画像 : 色・テクスチャ・輝度勾配 etc..
- 音声 : 周波数解析
- テキスト : 似た意味の言葉を集計したり



対象のデータをよく観察して
設計するしかない (もはや職人芸)

結局人間が大きく介在している

2. 特徴量に対する学習・推論を行うモデルの学習

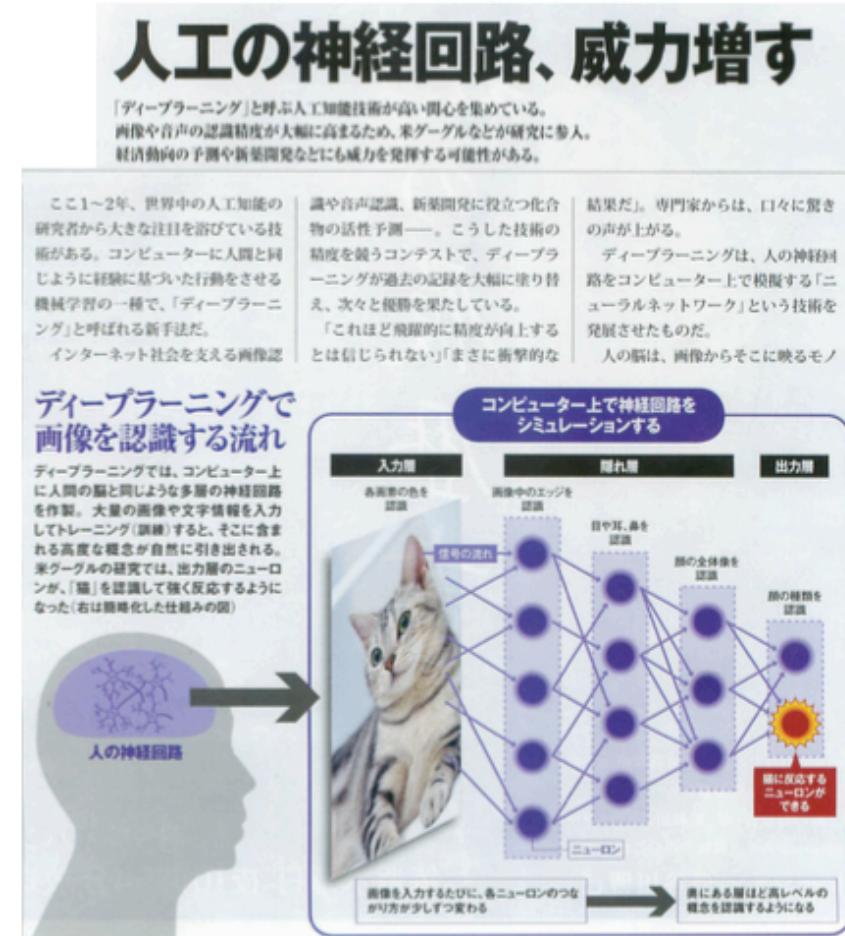
- 線形分離器 (SVM, Logistic回帰)
- 非線形分離器 (決定木, Kernel法, Bayes学習)
- サンプルベース (kNN)

Deep Learning

- ・ 松尾先生曰く「AIにおける50年来のブレイクスルー」

The screenshot shows the MIT Technology Review website with the header "MIT Technology Review". Below it, a navigation bar includes "LISTS", "INNOVATORS UNDER 35", "DISRUPTIVE COMPANIES", "BREAKTHROUGH TECHNOLOGIES", and "NOMINATIONS". The main content is titled "10 BREAKTHROUGH TECHNOLOGIES 2013". The "Deep Learning" entry is highlighted with a red border. Other technologies listed include Temporary Social Media, Prenatal DNA Sequencing, Additive Manufacturing, Baxter: The Blue-Collar Robot, Memory Implants, Smart Watches, Ultra-Efficient Solar Power, Big Data from Cheap Phones, and Supergrids.

Deep Learning	Temporary Social Media	Prenatal DNA Sequencing	Additive Manufacturing	Baxter: The Blue-Collar Robot
With massive amounts of computational power, machines can now recognize objects and translate speech in real time. Artificial intelligence is finally getting smart.	Messages that quickly self-destruct could enhance the privacy of online communications and make people freer to be spontaneous.	Reading the DNA of fetuses will be the next frontier of the genetic revolution. But do you really want to know about the genetic problems or musical aptitude of your unborn child?	Skeptical about 3-D printing? GE, the world's largest manufacturer, is on the verge of using the technology to make jet parts.	Rodney Brooks's newest invention is easy to interact with, but the complex innovations behind the robot show just how hard it is to get along with people.
Read more	Read more	Read more	Read more	Read more



Deep Learning

様々な分野で圧倒的性能

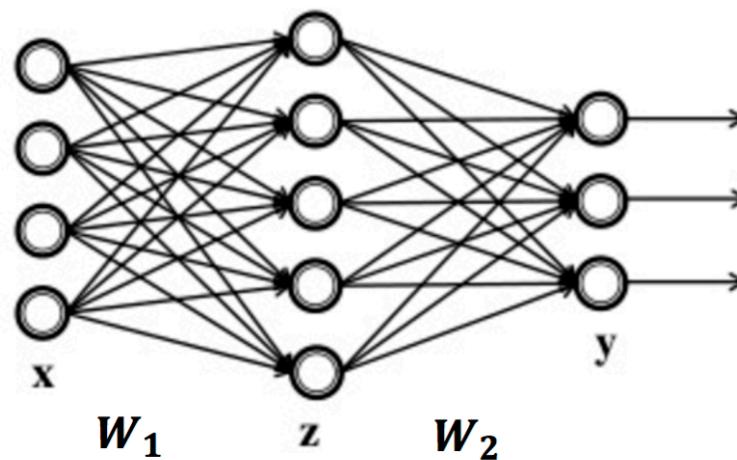
- 画像認識
 - ILSVRC 2012 では二位以下にerror rate 10%以上の差をつけ圧勝
- 音声認識
 - 従来の手法より5%以上認識率の向上
- 化合物の活性予測
 - 非専門家がぶつちぎりの1位
 - 「化合物の活性予測なんてやったことないが、試してみたら精度が出た。」とのこと
- 仏英翻訳タスク
 - 一切前処理なしでその時点での最高精度に匹敵

Deep Learningとは.. の前に

- ニューラルネットワーク
 - 提案自体は1960s ~
 - 脳の構造(シナプス)を参考に作成されたとされている
 - 基本的には、入力層、中間層、出力層の3層からなる

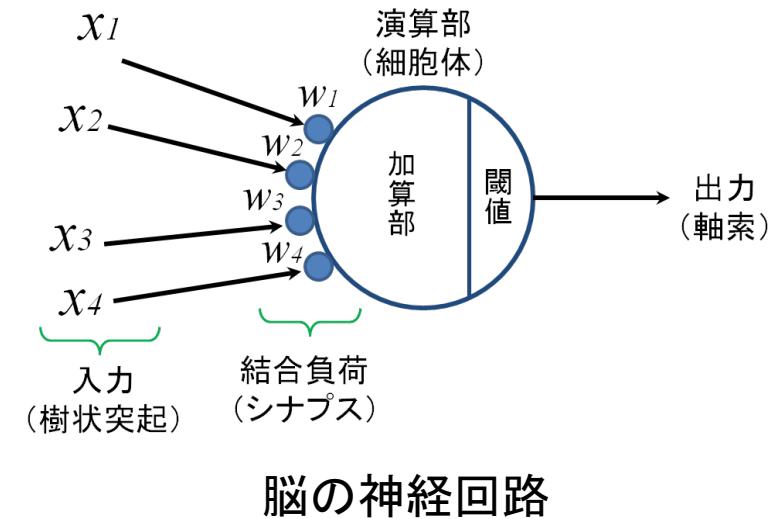
入力層 中間層(隠れ層) 出力層

3



$$y = f(W_2 \cdot f(W_1 \cdot X) + b) + b$$

(人工)ニューラルネットワーク

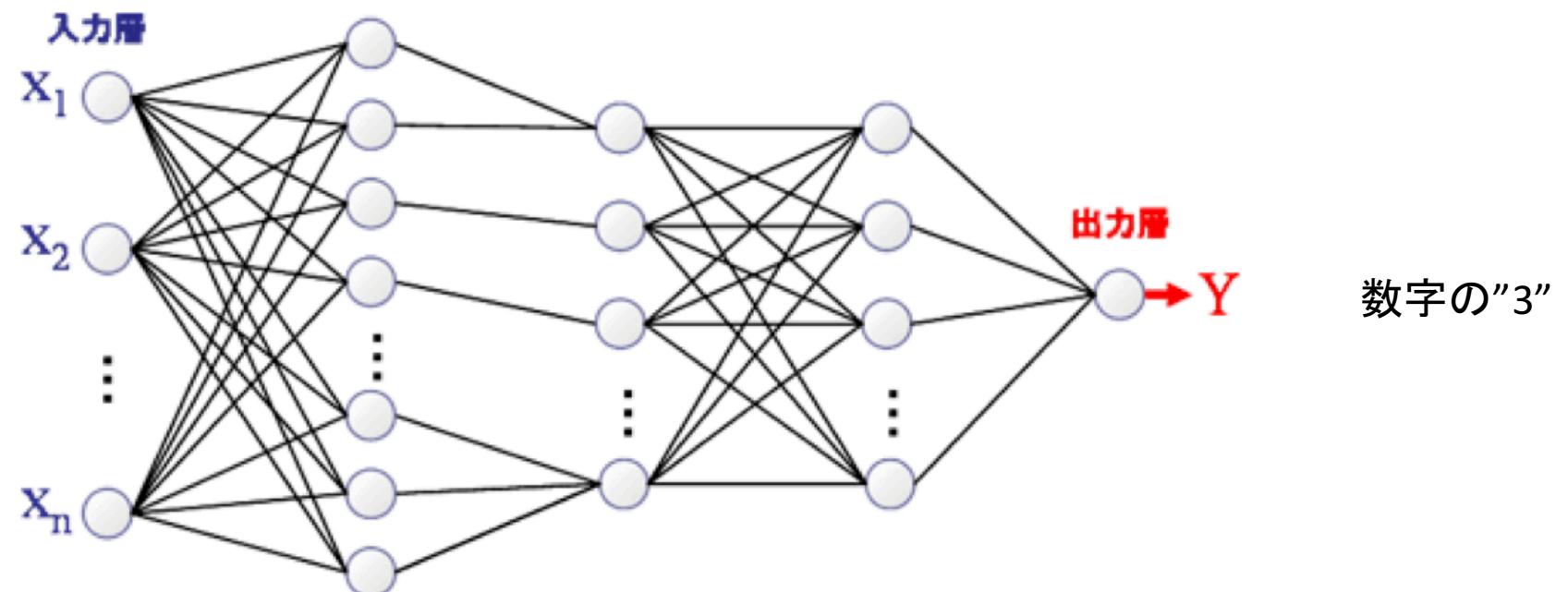


脳の神経回路

Deep Learningとは

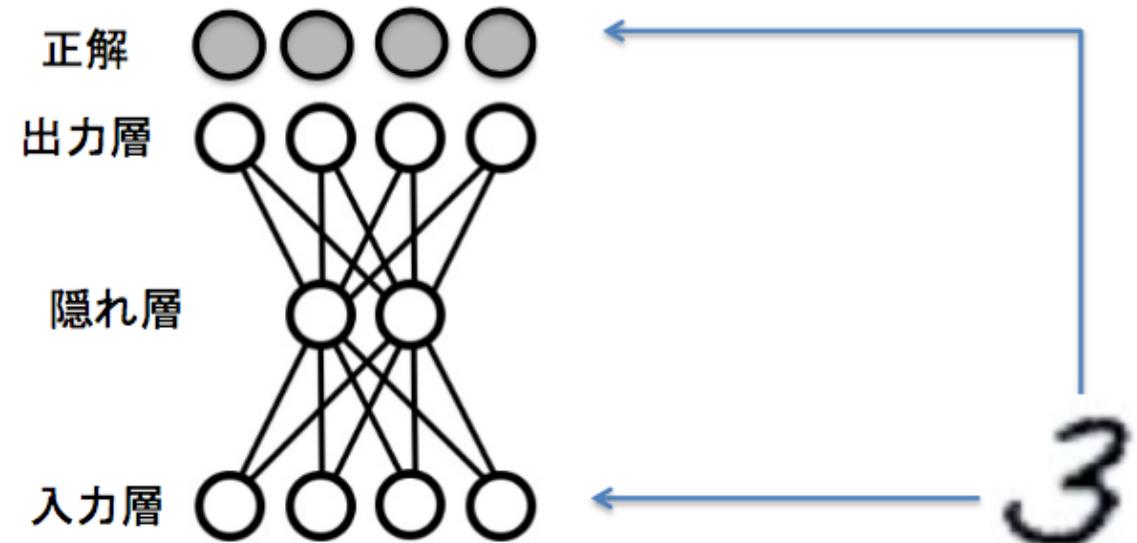
ニューラルネットワークを多層に(深く)積み上げた
ディープニューラルネットワークを利用した機械学習技術の総称

3



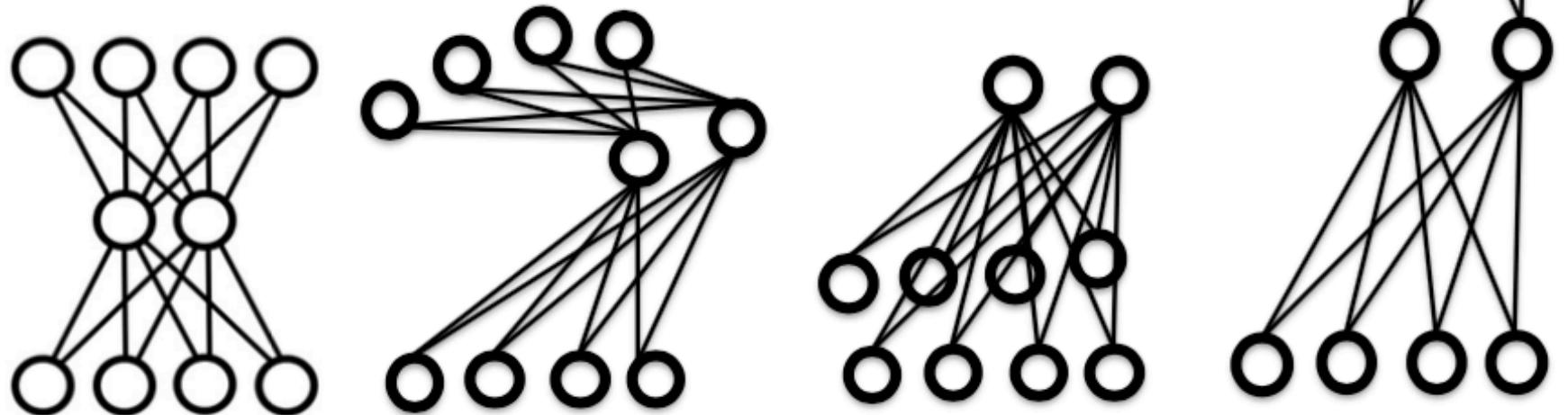
Auto Encoder

- Deep Learningが最初に注目されるきっかけ(正確には再注目)
- 出力と入力を同じにして学習するニューラルネットワーク
- 分類としては教師なし学習
- 普通に考えると意味ないが、隠れ層の次元数が入力よりも小さかった場合、より小さい情報で元の情報を表現(情報圧縮)



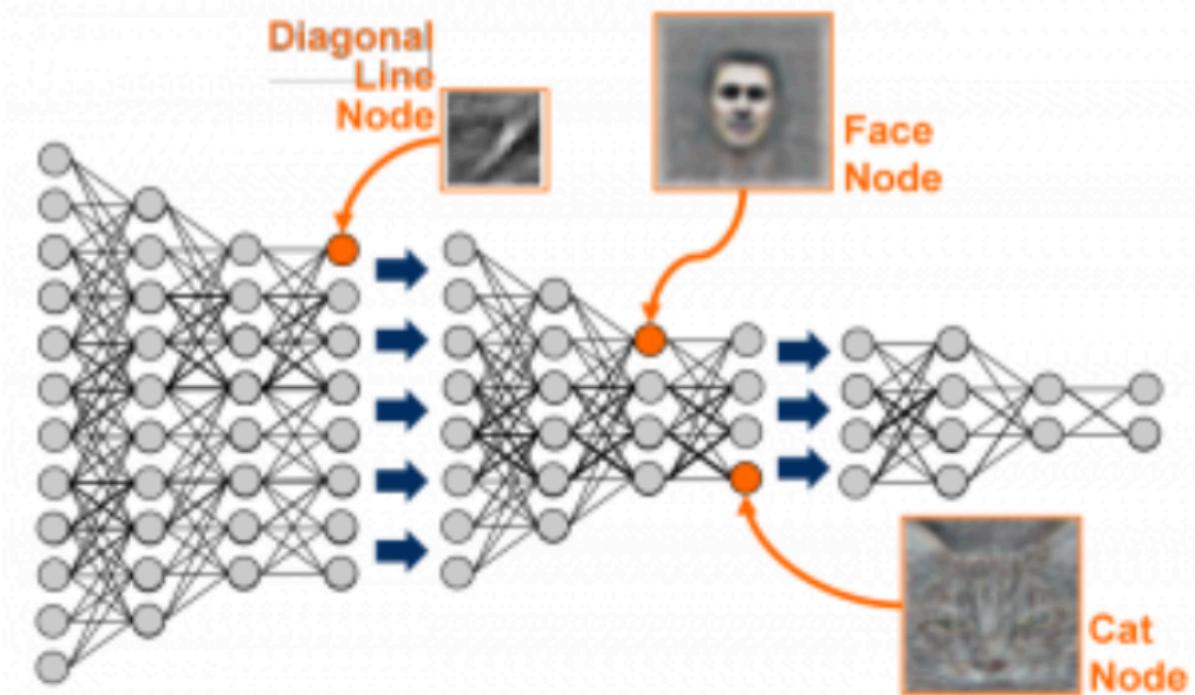
Deep AutoEncoder

- 層を重ねることで、どんどん小さい情報で表現しようとする
- 最終的に各ユニットが何らかの幾何情報を表現



Googleの猫

- Googleの研究
- Youtubeから取得した大量の画像を使って先程のAutoEncoderを学習
(若干違うがコンセプトは同じ)
- 下位の層のニューロンではななめの線とかエッジに反応するように
- 上位では人の顔や猫といった複雑な特徴に反応するように階層的な表現を獲得
- 人間の視神経のモデルとして知られるものと近いらしい



何がすごいのか

機械学習の2ステップ

1. 入力データから人手で特徴量を設計、抽出する
2. 特徴量に対する学習・推論を行うモデルの学習

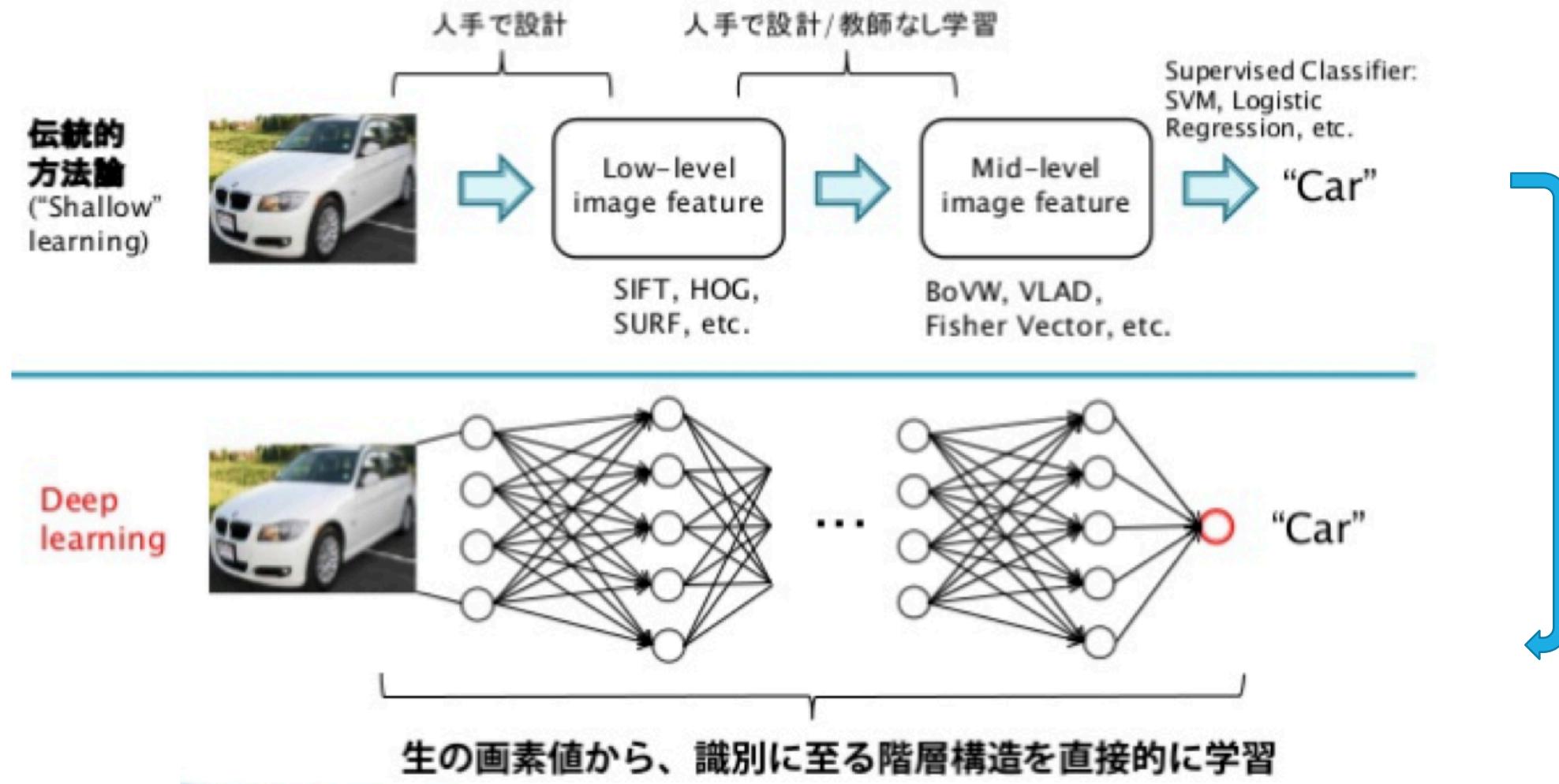
表現学習と呼ばれる

ステップ1, 2を**同時に**行うことが出来る
データから自動的に特徴量を抽出することができる。

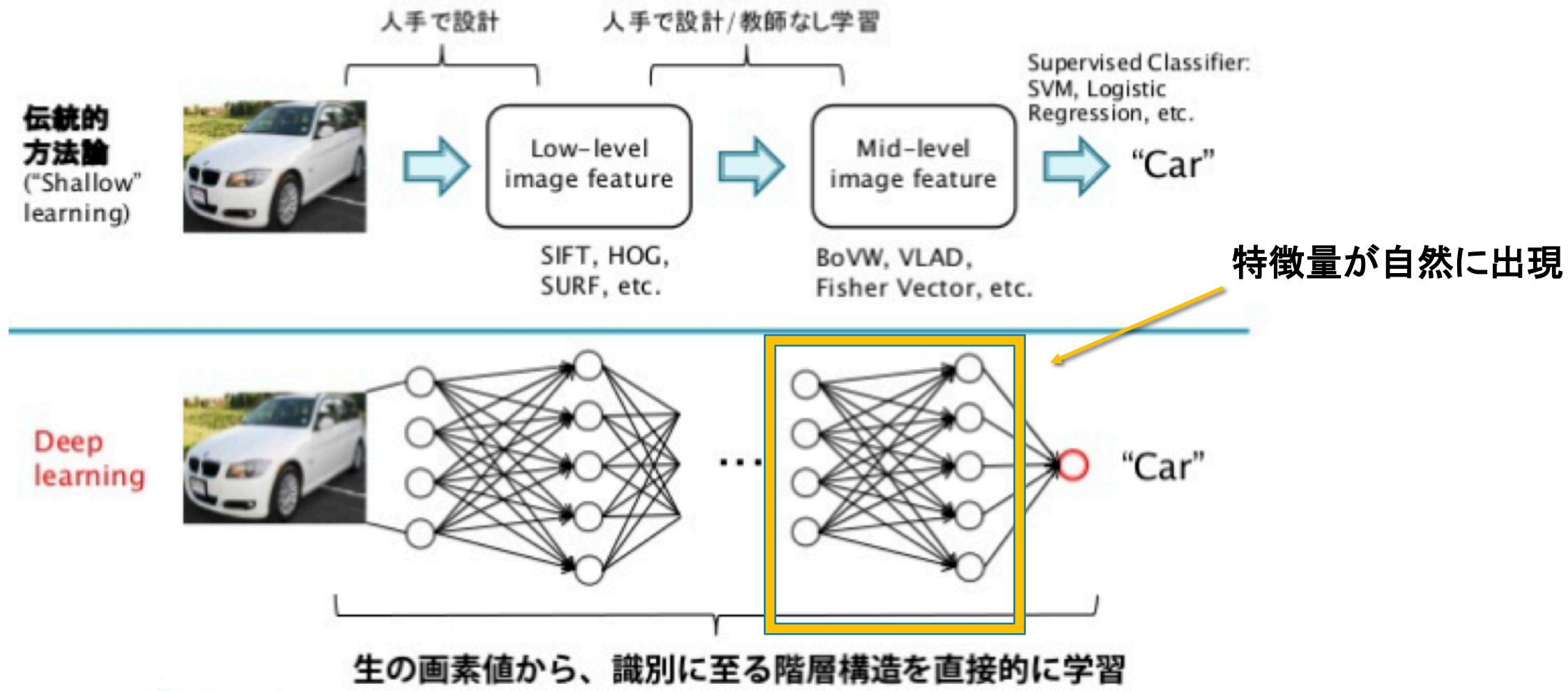


そのうえ人手で設計した特徴量よりも**性能が良い**
しかも、**異なる分野**で同様に適用できた

例. 画像認識パイプライン



例. 画像認識パイプライン



画像認識

- ILSVRC2012 : Large Scale Visual Recognition Challenge 2012



Team name	Error	Description
SuperVision	15.315%	Using extra training data from ImageNet Fall 2011 release
SuperVision	16.422%	Using only supplied training data
ISI	26.602%	「ケタ」が違う scores from classifiers using each FC
ISI	26.646%	Naïve sum of scores from classifiers using each FV
ISI	26.952%	Naïve sum of scores from each classifier with SIFT+FV, LBP+FV, GIST+FV and CSIFT+FV, respectively
OXFORD_VGG	26.979%	Mixed selection from High-Level SVM scores and Baseline Scores, decision is performed by looking at the validation performance.
...

ディープ
ラーニング

長年の
特徴量設計
の工夫

画像認識

- ILSVRC2012 : Large Scale Visual Recognition Challenge 2012



常に引き合いに出される
2位のISIというのは実は
原田・国吉研究室

Team name	Score	Description
SuperVision	16.422%	Using only supplied training data
ISI	26.602%	「ケタ」が違う scores from classifiers using each FC
ISI	26.646%	Naïve sum of scores from classifiers using each FV
ISI	26.952%	Naïve sum of scores from each classifier with SIFT+FV, LBP+FV, GIST+FV and CSIFT+FV, respectively
OXFORD_VGG	26.979%	Mixed selection from High-Level SVM scores and Baseline Scores, decision is performed by looking at the validation performance.
...

ディープ
ラーニング

長年の
特徴量設計
の工夫

エラー率 変化:2012年以降

	Error
Before ディープ ラーニング	Imagenet 2011 winner (not CNN) 25.7%
After ディープ ラーニング	Imagenet 2012 winner 16.4% (Krizhevsky et al.)
	Imagenet 2013 winner 11.7% (Zeiler/Clarifai)
	Imagenet 2014 winner 6.7% (GoogLeNet)
	Baidu Arxiv paper: 2015/1/3 6.0%
	Human: Andrej Karpathy 5.1%
	Microsoft Research Arxiv paper: 2015/2/6 4.9%
	Google Arxiv paper: 2015/3/2 4.8%
	Microsoft Research CVPR paper: 2015/12/10 3.6%

2015年2月には人間の精度を超えた

画像認識で人間の精度を超えることは
数十年間、実現されていなかった⁴⁰

顔認識

- Googleの研究
- 2億枚の人間の顔画像から学習
- 精度 : 99.63% !!!

→は別人なのに機械が同一人物だとした例



・人間が見ても判定が難しい

F. Schroff et. al: FaceNet: A Unified Embedding for Face Recognition and Clustering, 2015

顔認識

- Googleの研究
- 2億枚の人間の顔画像から学習
- 精度 : 99.63% !!!

→は同一人物なのに機械が別人だとした例



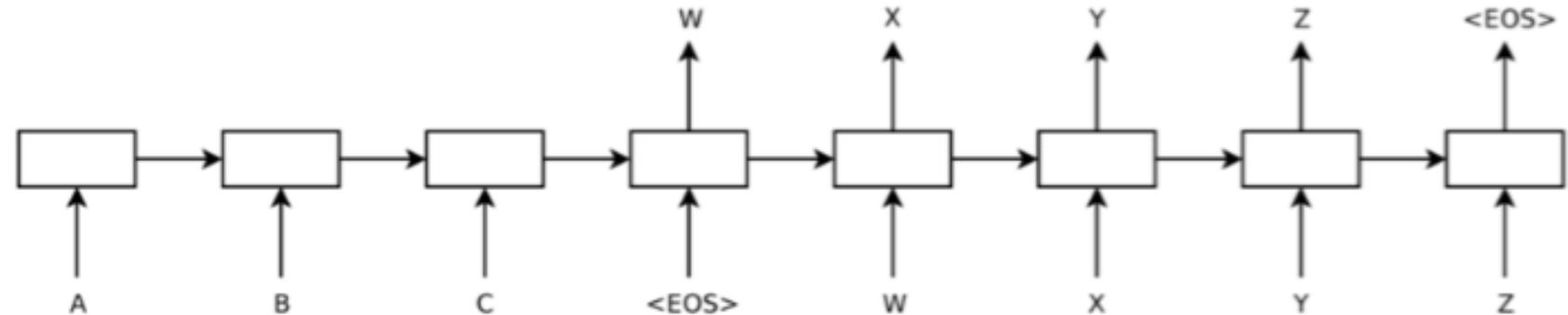
・正解データの信ぴょう性のほうが問われる。

F. Schroff et. al: FaceNet: A Unified Embedding for Face Recognition and Clustering, 2015

翻訳

- 状態を保存するようなニューラルネットワーク
- 片方の言語を入力したら、同じ意味のもう片方の言語が出力されるように学習
- 最近Google翻訳が劇的に性能が上がったのはこれを採用してから

Machine Translation [Sutskever et al. 2014]



最近の研究

- 大量のデータと正解が与えられた場合にそれらの対応 (マッピング)を学習する問題に関してはほぼなんでもできてしまうように
 - データが少量時の学習、高速な学習、省メモリな学習、etc...
- 複数のモダリティを扱う研究 (マルチモーダルな学習)が比較的盛ん(モダリティ: 画像・テキスト・音声・属性など)
 - 画像とその説明文が与えられて、未知の画像に対してもそのキャプションを生成
 - その逆で、説明文が与えられて画像を生成
(情報量に大きな差があるのでさすがにまだ微妙だが)
 - 動画から音声を類推する

Image to Caption



"man in black shirt is playing guitar."



"construction worker in orange safety vest is working on road."



"two young girls are playing with lego toy."



"boy is doing backflip on wakeboard."



"girl in pink dress is jumping in air."



"black and white dog jumps over bar."



"young girl in pink shirt is swinging on swing."



"man in blue wetsuit is surfing on wave."

Caption to Image



A very large commercial plane flying in blue skies.



A very large commercial plane flying in rainy skies.



A herd of elephants walking across a dry grass field.



A herd of elephants walking across a green grass field.

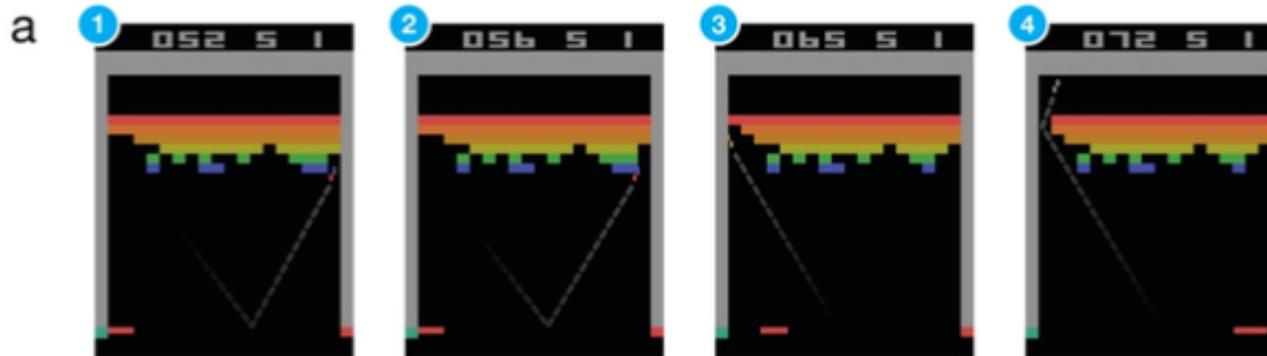
強化学習 × Deep Learning

強化学習

- 機械学習の枠組みの中でも若干特殊で、
学習の主体が自ら試行錯誤して学習していく
 - 「行動」に対して、「報酬」が得られると、直前の行動を評価する
 - ある「状態」のときにどういう「行動」を取れば、得られる「報酬」が最大になるようになるのか学習する
 - 動物が恐らく生来的に行っている学習
- 古くからある技術だが、今まででは状態の定義が人間によるものだったので、
簡単な問題しか解けなかつた
 - 迷路
 - バックギャモンなどの簡単なゲーム

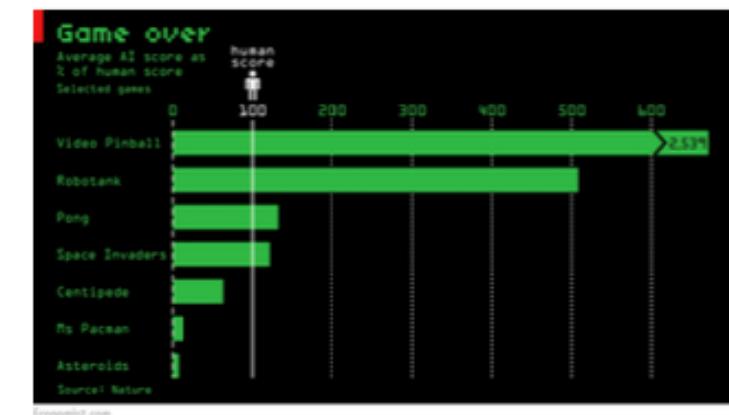
強化学習 × Deep Learning

- ビデオゲームの学習
- 状態をニューラルネットワークで表現
- DeepMindの研究(現在はGoogleに買収)
- 試行錯誤により、運動(ゲームスキル)が習熟
 - 最初は何もできない。が、繰り返すうちに上手に
 - 画像中のどこを見ればいいのかを自動で学習
 - 最終的には人間を上回るスコアを獲得



紛らわしいところ

- 試行錯誤により自律的に学習していく
枠組みが**強化学習**
 - 画像から特徴量をいい感じに抽出しながら
学習するのが**Deep Learning**
- ほとんどの記事はこのへんごっちゃにして書いてます!



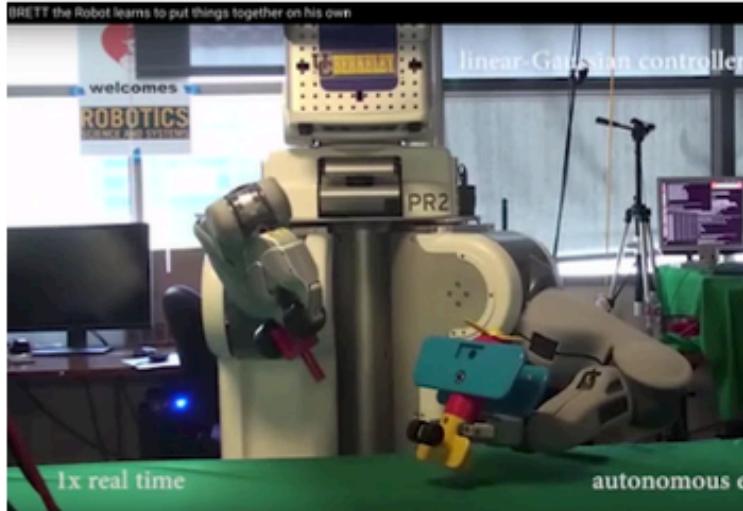
強化学習 × Deep Learning

- 実世界への適用
 - 部品の取り付けを習熟するロボット (UC berkeley)
 - 運転を習熟するミニカー (PFN)
 - ピッキングを習熟するロボット (PFN・ファナック)

<https://www.youtube.com/watch?v=7A9UwxvgcV0>

https://www.youtube.com/watch?time_continue=1&v=ATXJ5dzOcDw

試行錯誤で作業学ぶロボット (UC Berkeley)



試行錯誤で運転を学習するミニカー (PFN社、日本)



ピッキングを習熟するロボット (PFN・ファナック)



Deep Learningの登場によって

- 生データからどこに注目すれば良いのかを自動的に判別してくれるよう
 - 確かに人間は、画像の輝度勾配が物体の識別に有効なんて教わったことはない
 - 人間に近い学習法
- 特徴設計が面倒、困難だったことが原因で出来なかつたことができるよう！！
 - ゲームがとてもいい例
 - また、センサー情報などどう扱えばいいかよくわからない物もよしなにやってくれる
- モラベックのパラドックス：「子供のできることほど難しい」
 - 高度な推論や、記憶処理よりも、認識や運動スキルのほうがよっぽど難しい
 - しかし、ここ最近の研究で兆しが見えてきた。

Deep Learningの登場によって

しかし、

- 本質的に新しいことができるようになったのか？
 - 一部例外はあるが、大半はそういうわけではない
 - 実用的に使用できる精度にまで向上したという表現が正しい
- Deep Learningはなんでも出来るのか？ 従来手法はもう使わないのであるのか？
 - 確かに、画像、音声、テキスト等の特徴抽出が極めて難しい分野においては劇的な性能向上が見られたが、テーブルデータ(年齢、性別等)やでは従来手法で十分な性能が出ることが多い
 - また、データから学習するということは、それだけ大量に答えも要るのでデータ作成コストが高い
 - さらに、膨大な計算コストがかかるので、それまた難しい場合も
 - 特に画像系は処理が重く、数台～数10台の高性能パソコンを数週間稼働させることもザラ

Deep Learningすごい！Deep Learning以外は古い！というのは危険

第3次AIブーム:機械学習から表現学習へ

- ビッグデータ・計算機能力・機械学習アルゴリズムの発展によって実用的なパターン認識が可能に
- Deep Learningによって従来の課題であった人手による特徴量設計を表現学習によって代替可能になった
- 複数の知覚にまたがった処理を扱うマルチモーダルな研究が盛ん
- 試行錯誤をしながら学んでいく強化学習とDeep Learningを組み合わせて運動の習熟を試みる研究も進んでいる

Agenda

1. 人工知能とは？
2. 人工知能技術の歴史
3. 近年の人工知能技術 (主に機械学習) の基礎
4. 産業へのインパクト

松尾先生が考える今後

ディープラーニングの今後の発展

① 画像

画像から、特微量を抽出する

画像認識の精度向上

② マルチモーダル

映像、センサーなどのマルチモーダルなデータから
特微量を抽出し、モデル化する

動画の認識精度の向上、行動予測、異常検知

③ ロボティクス(行動)

自分の行動と観測のデータをセットにして、特微量を抽出する。
記号を操作し、行動計画を作る。

プランニング、推論

④ インタラクション

外界と試行錯誤することで、外界の特微量を引き出す

オントロジー、高度な状況の認識

⑤ 言葉とのひもづけ(シンボルグラウンディング)

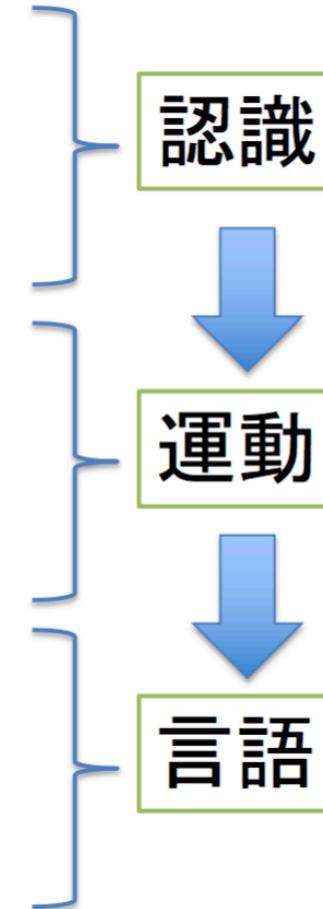
高次特微量を、言語とひもづける

言語理解、自動翻訳

⑥ 言語からの知識獲得

グラウンディングされた言語データの大量の入力により、さらなる抽象化を行う

知識獲得のボトルネックの解決

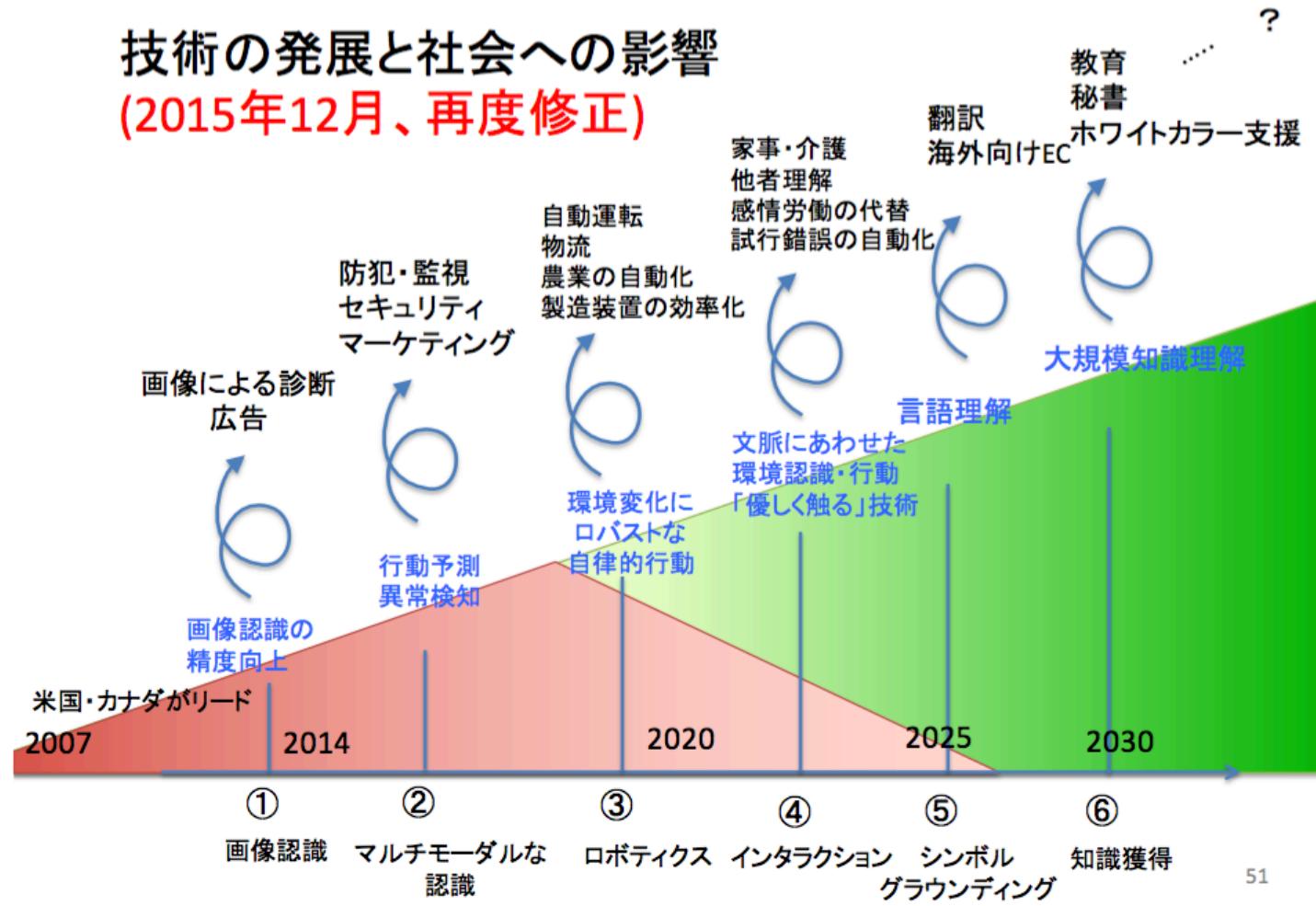


松尾先生が考える今後

- 認識(画像・映像)
 - 画像認識が出来ないから人間がやっている仕事がたくさん
→ ほぼ機械によって代替可能
Ex. 監視などにかかるコストは100分の1以下に

- 運動の習熟
 - 機械は「機械的」な動きしかできないから単純だが、柔軟な動きが必要なタスクは人間が行っている
→ 機械が「機械的でない」動きを可能に
Ex. 複雑な形状の物体のピッキング作業、食品加工

松尾先生が考える今後



松尾先生が考える今後

既存産業の発展

農業	収穫判定	トラクター、コンバインの適用範囲拡大、効率向上 選別調製等の自動化	自動での収穫 自動での耕耘
建設	測量	掘削、基礎工事、外装内装作業等の効率向上	多くの作業の自動化・効率化
食品加工	振り分け確認	カット、皮むき、解体等の自動化	多くの加工工程の自動化
組み立て加工	目視確認の自動化	動作効率の向上	段取りの自動化 セル生産の自動化

⋮

A: 画像認識

B: 運動の習熟

C: 計画立案を伴う運動

松尾先生が考える今後

日本の社会課題と人工知能

- 農業分野に「習熟したロボット」を適用することで
 - 休耕地が耕せる。除草・防除や収穫ができる。収量が増える。
- 介護分野に適用することで
 - 介助も楽に。移動したりトイレにいけるようになり、より自立した生活ができる。
- 廃炉作業に適用することで
 - 危険な状況で人が作業しなくてよくなる。工期を短縮できる。
- 河川や火山を見張ることで
 - 河川の氾濫や土砂崩れ、噴火などの危険な状態・予兆を早期に発見できる。
- こうした技術を使った製品を海外に展開していくことで
 - 新たな輸出産業に。GDPの増加につながる。

終わりに

- 世間ではシンギュラリティという言葉もしばしば
 - シンギュラリティ(技術的特異点)
 - 人工知能に対するかなり極端な観点
 - AIが自分より賢いAIを作れるようになった時、一気に発散するという話 ($1.01 * 1000 \rightarrow \infty$)
 - 周りの人工知能研究者で、これがすぐに訪れると考えてる人は誰もいない

→ 人間 = 知能(考える力) + 生命(目的意識)だとすると、

→ 結局、どこまで言っても目的は人間から与えられるだけ

→もちろん悪意を持った人間が悪用すれば危険(そんなの車でも包丁でも同じ)

終わりに

- 「人工知能に仕事を奪われる」
 - 当然なくなる単純作業労働は増えるし、働き方が変わる仕事もある
 - が、ミシンの出現で裁縫仕事を奪われたのと別に何ら変わらない
 - 単純作業の自動化によって、人間にしか出来ないことの重要性は増していく
- しっかり動向、性質を追いその上で社会、仕事に活かすことが肝要
 - 世間のニュースなどに振り回されない(基本的に間違っています)

まとめ

1. 人工知能とは？

- 汎用人工知能、特化型の2種類。ほぼ後者

2. 人工知能技術の歴史

- 第1次AIブーム: 探索・推論の時代 (1956~1960年代)
- 第2次AIブーム: 知識処理の時代 (1980年代)
- 第3次AIブーム : 機械学習の時代 (2000年代~)

3. 近年の人工知能技術(主に機械学習)の基礎

- 計算機能力の向上とアルゴリズムの改善で機械学習が発展
- Deep Learningに登場により人手による特徴量設計のボトルネックを解決

4. 産業へのインパクト

- 様々な応用可能性 特に画像認識、ロボット