

Deep Learning 輪読会 2017

第1章 はじめに

2017.11.06

東京大学 松尾研究室

リサーチエンジニア 曾根岡侑也

構成

1章 はじめに

1.1 誰がこの本を読むべきか

1.2 深層学習の動向の歴史

1.2.1 ニューラルネットワークの数多くの名前と変わりゆく運命

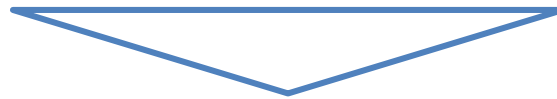
1.2.2 増大するデータ集合のサイズ

1.2.3 増大するモデルサイズ

1.2.4 精度, 複雑生, 実世界の影響の増大

Artificial Intelligenceの歴史（概観）

- 古代ギリシャ時代から、「考える機械」を発明家は夢見てきた
- 1842年に**コンピュータが考案**され、「知的なもの」に対する議論がなされた
- AIの初期には、人間にとっては難解だが**コンピュータに簡単なタスク**に取り組まれるようになっていった（例：チェス：ルールの記述が簡単）
- コンピュータへの記述が難しいが**人間には直感的で簡単なタスク**へ（例：話し言葉や画像の認識）



DeepLearning：直感的な問題を解くためのアプローチ

Artificial Intelligenceの歴史：コンピュータで解くべきタスク

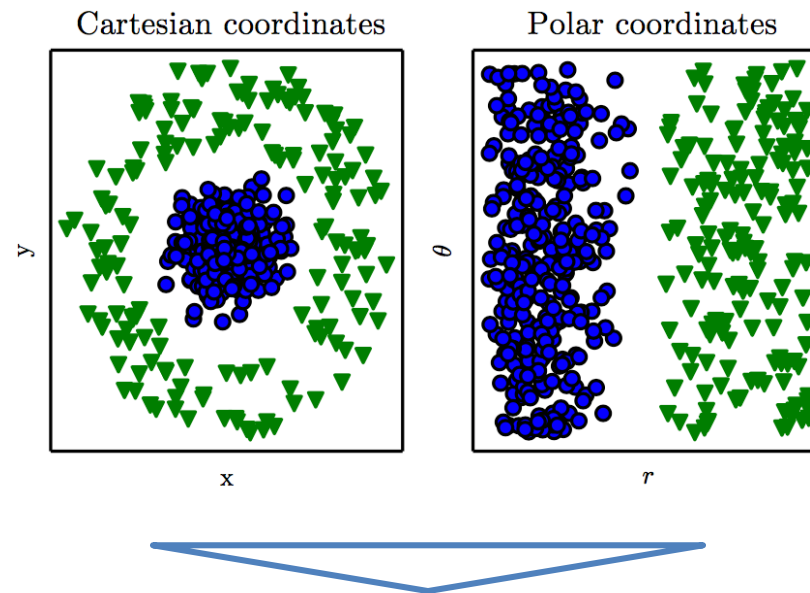
- **定型的な環境のタスク** → 初期に解決
 - (例) チェス：64マスと動かし方が決まった32コマで定義される
 - 1997年にIBMのDeepBlueが世界チャンピオンを倒す
- **型にはまらないタスク** → 人間には簡単だが、コンピュータでは難しい
 - (例) 物体や音声の認識
 - 途方もない量の主観的かつ直感的な知識が必要
 - 人工知能でずっと取り組まれてきた課題

Artificial Intelligenceの歴史：型にはまらないタスクへのアプローチ

- 型にはまらないタスクを解決するために数々のアプローチが取られた
- **アプローチ①：知識ベース**
 - 形式言語で全ての知識をハードコーディングする方法
 - 例：推論エンジンCycと形式言語CycLで書かれたデータベース
 - 問題点：**世界を正確に記述する形式的なルールを考案できない**
- **アプローチ②：機械学習**
 - 生データからパターンを抽出し知識を獲得する方法（①の問題を解決）
 - 例：ロジスティック回帰・ナイーブベイズ
 - 問題点：**性能が与えられたデータの表現（特徴量）に依存**
（どういう情報を入力にするかは人次第）

表現に対する依存性

- データの表現次第で情報処理タスクの難易度は変わる
 - 例1) アラビア数字の暗算 (「CCX割るVI」 vs 「210 割る 6」)
 - 例2) 数値リストのソート (連結リスト vs 木)
 - 例3) 下図のデータを直線で分離 (デカルト座標 vs 極座標)



機械学習では表現（特徴量の設計）が重要であるが、
正しい特徴量が何かを知ることは難しい

Artificial Intelligenceの歴史：型にはまらないタスクへのアプローチ

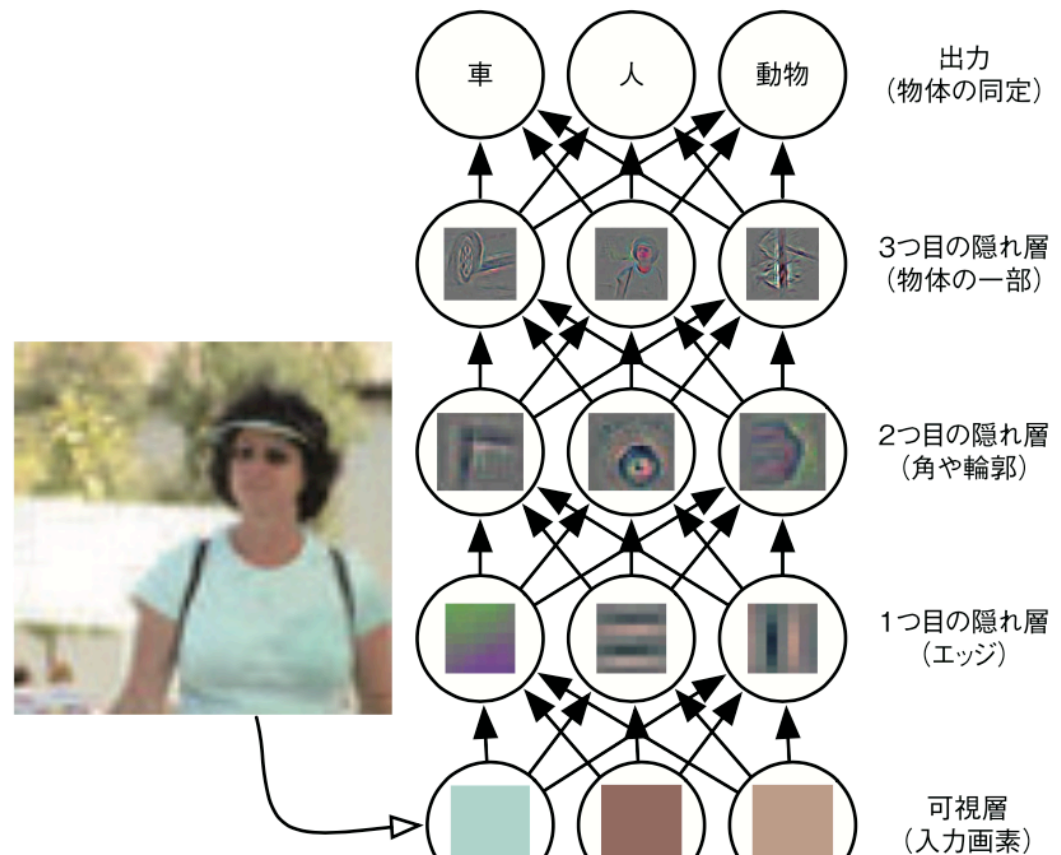
- アプローチ③：表現学習

- 表現から出力の写像だけではなく表現自体も学習する（②の問題を解決）
- 人の介入を最小限に押さえて新しいタスクにすぐ適応できるようになる
- 例：自己符号化器（Autoencoder）
 - 入力データを異なる表現に変換する符号化器（Encoder）と
表現を元の形式に戻す復号化器（Decoder）の組み合わせ
- 問題点：観測できる量に影響を与える要因は複雑かつ多用で、
要因のもつれを解いて、重要でないものを取り除く必要がある

Artificial Intelligenceの歴史：型にはまらないタスクへのアプローチ

• アプローチ④：DeepLearning

- 階層的に単純な表現から複雑な表現を構築する手法（③の問題を解決）
- エッジから輪郭などのパーツ、物体の一部という階層で表現を獲得
- （例）多層パーセプトロン

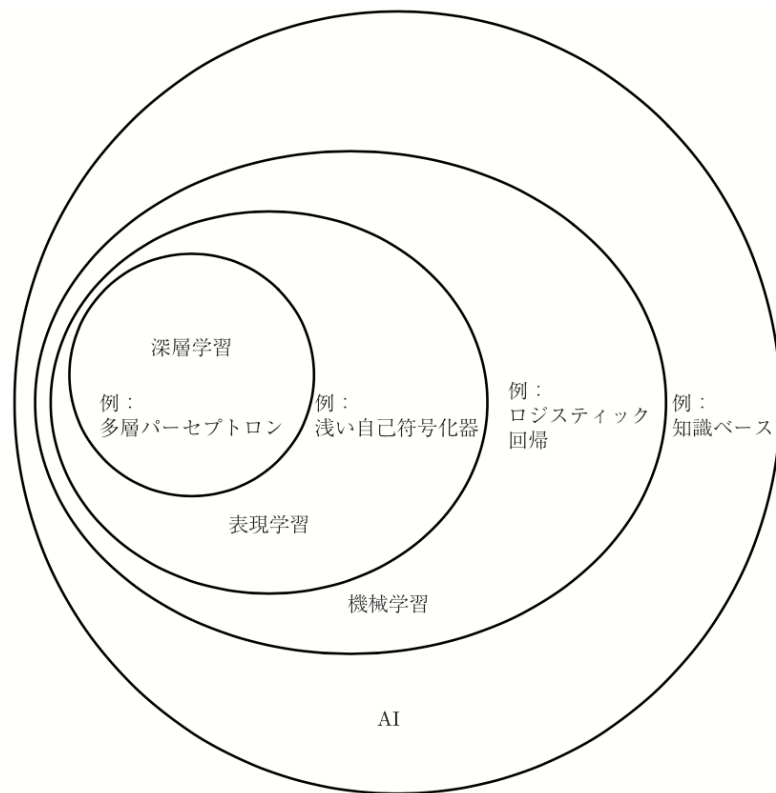


深さの定義

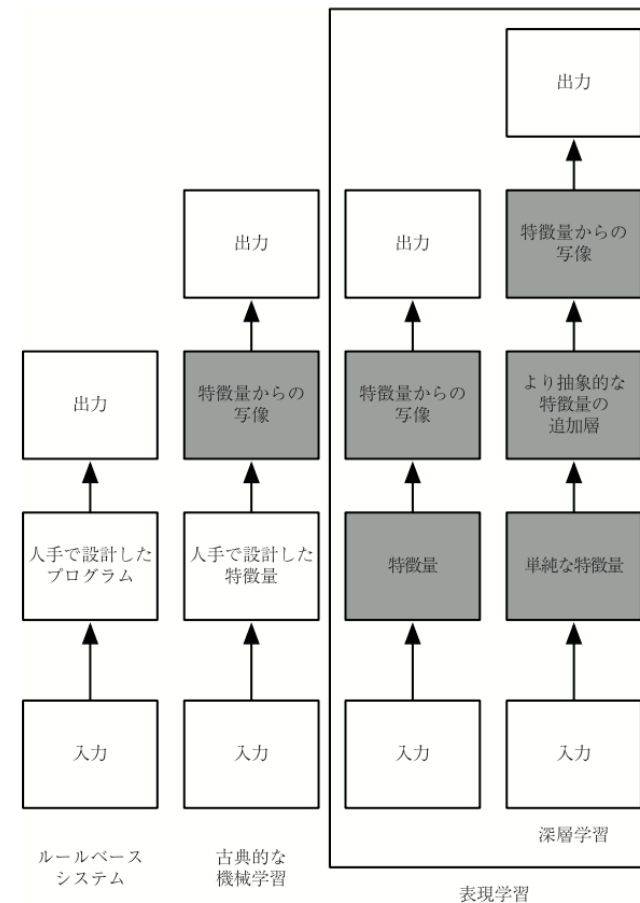
- DeepLearningはどれくらいの深さだと「Deep」になるのか？
- 深さの定義
 1. 計算グラフの深さ
 - 出力を計算する上でのフローチャートの長さ
 - 関数の要素の切り方で長さが変わる
 2. 確率モデルのグラフの深さ
 - 概念の階層の深さ
- 深さ自体も一意に決まらず、**「モデルが深い」とされる深さの合意はない**

DeepLearningの立ち位置

- DeepLearningは機械学習、さらには表現学習の一分野であり、階層的により抽象的な特徴量を獲得する点が従来手法との差分



DeepLearningの立ち位置



フローチャート

1.1 誰がこの本を読むべきか

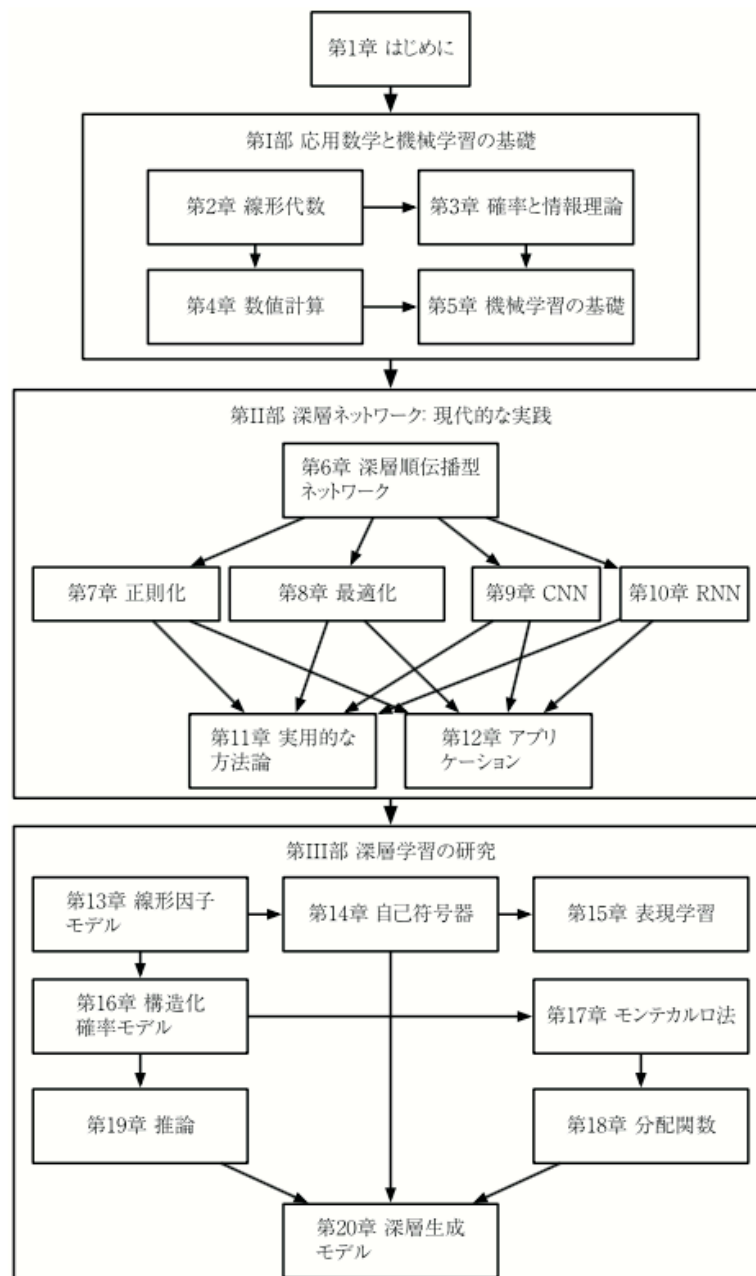
- **想定読者**

- 機械学習を学ぶ大学の学生(学部生または大学院生)
- 深層学習や人工知能研究のキャリアを始めようと考えている人々
- 機械学習や統計学に関する経歴がないソフトウェアエンジニア

- **事前知識**

- プログラミング、計算性能の問題、微積分、グラフ理論に関するある程度の知識

1.1 本書の構成



基礎的な数学

・・・既に学習したものは飛ばしていい

既に確立された技術

将来の深層学習の研究で 重要だと信じられている分野

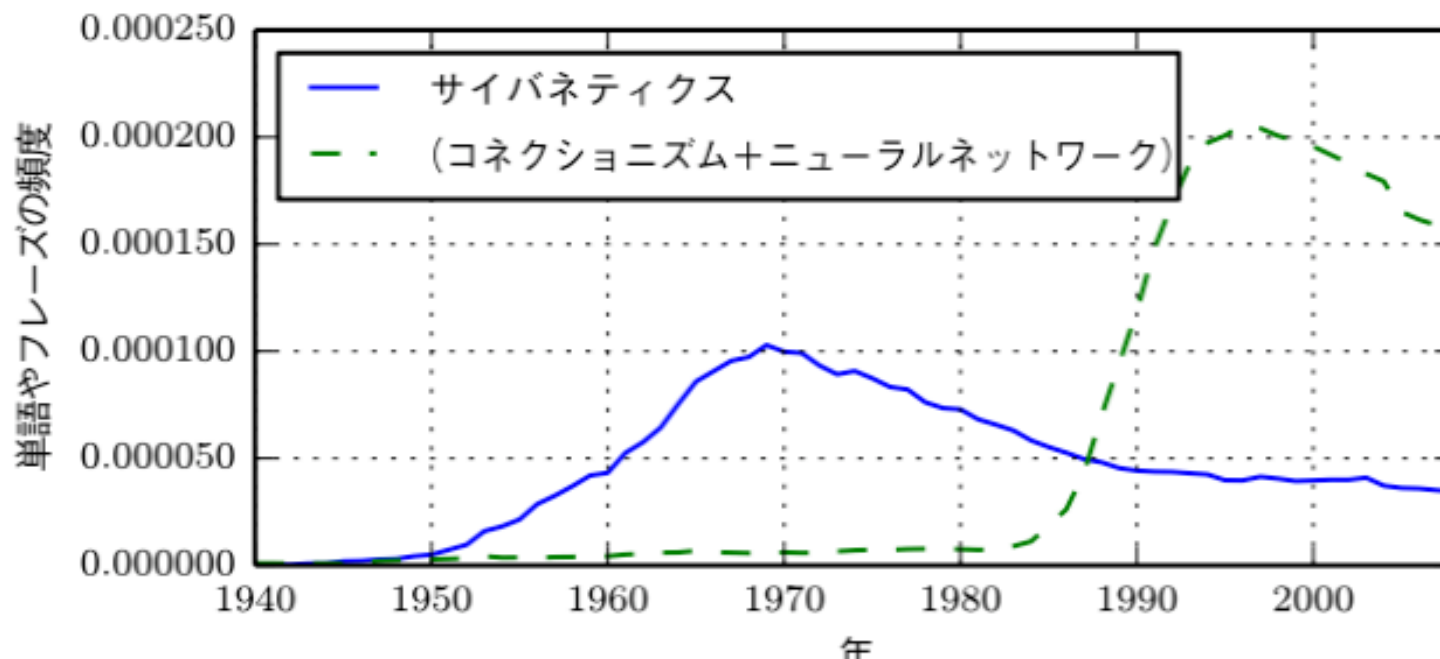
・・・現時点の社会実装に興味ある人は
飛ばしていい

1.2 深層学習の動向の歴史

- 深層学習の主要な動向は下記
 - (1.2.1) 歴史は長く今まで様々な名前がつけられ人気の盛衰があった
 - (1.2.2) 利用可能な学習データの量が増え、有用性が高まった
 - (1.2.3) コンピュータ基盤の改善により、モデルサイズが大きくなった
 - (1.2.4) 精度向上に伴い、応用分野が広がった

1.2.1 ニューラルネットワークの数多くの名前と変わりゆく運命

- 1940-1960年代：サイバネティクス
 - 生物学的学習理論とパーセプトロンの登場
- 1980-1995年：コネクショニズム or ニューラルネットワーク
 - 逆伝搬によるニューラルネットの利用
- 2006-現在：深層学習



1.2.1 歴史：深層学習の起源

- 脳のなかで学習がどのように起こるかというモデルを意図して、
人工ニューラルネットワークが作られた
- 神経学的な研究動機
 - 脳は知的な行動ができる実例であり、
その計算原理のリバースエンジニアリングしたい
 - 脳と知能の原理を解き明かしたい

1.2.1 歴史：サイバネティクス

- 神経科学の影響を受けた単純な線形モデル

$$f(\boldsymbol{x}, \boldsymbol{w}) = x_1 w_1 + \cdots + x_n w_n \quad (\boldsymbol{x} : \text{入力}, \boldsymbol{w} : \text{重み})$$

- **McCulloch-Pittsニューロン** (1943年)
 - 重み w は人手で設計し、 $f(\boldsymbol{x}, \boldsymbol{w})$ で2つのカテゴリ識別を行なった
- **線形モデル**の登場
 - **パーセプトロン** (1958年)
 - カテゴリ識別（分類）で、入力事例から重みを学習するモデル
 - **ADALINE** (1960年)
 - 実数値の予測（回帰）で、入力事例から学習するモデル
 - **確率的勾配降下法**が登場
- **XORを学習できない**という問題でブームは下火に

1.2.1 歴史：コネクショニズム

- 神経機構に基づいた**認知モデル**の研究が発端（1985年）
 - 従来の記号モデルでは脳がニューロンを使ってどのようにモデルを実装しているかの説明が難しかった
 - 大量の単純な演算ユニットの集合で知的な振る舞いを再現するアイデア
- 数々の概念・アルゴリズムがこの時期に生まれた
 - **分散表現**（1986年, Hinton先生, 15章）
 - **誤差逆伝播法**（1986・1987年, LeCun先生）
 - **LSTM**（1997年）
- **カーネルマシン・グラフィカルモデルの活躍**で下火に

1.2.1 歴史：DeepLearning

- 1995年以降下火の中、一部の研究機関はニューラルネットの研究を進めた
 - CIFAR, Hinton先生, Bengio先生, LeCun先生ら
- 2006年：DeepBeliefNetによるブレイクスルー
 - 深いネットワークの学習は難しかった
 - 層ごとの**貪欲事前学習**で効率的に学習できることを示した
- 2007年～：少しずつ成果が出てきており、
10年たった今も関心がより集まっている

1.2.1 歴史：神経科学との関係

- **神経科学は重要な着想の源**

- 起源となったニューラルネットワークの着想に
ネオコグニトロン（CNN）, ReLU, モデルアーキテクチャの着想も
- 脳の中は単一のアルゴリズムでできている可能性を示唆
 - フェレットの脳の視覚と音声処理を行うところを交換しても学習できたという実験結果
 - 深層学習で画像・音声・NLPが一緒に研究されている

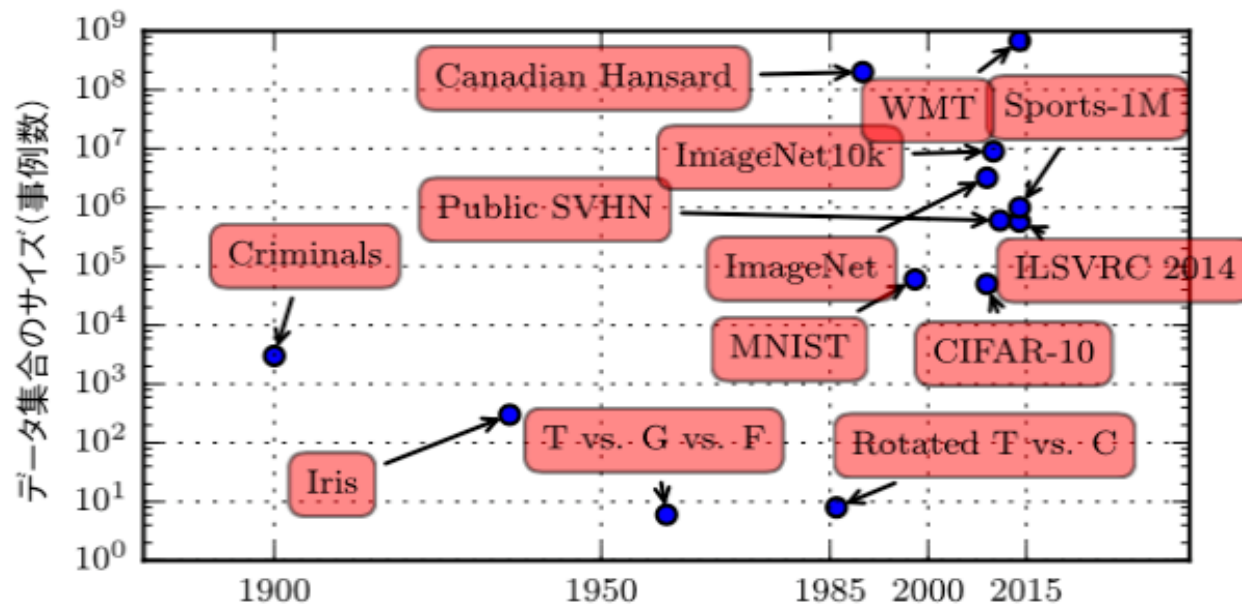
- 一方で、神経科学は**主要な案内役ではなくなっている**

- 脳についての情報が十分ではない
- 最も単純で十分に研究されているところでさえ理解からは程遠い

- 脳から着想を得ているだけの**深層学習**に対して、
脳の正確なモデルを解明しようとする**計算論的神経科学**が存在

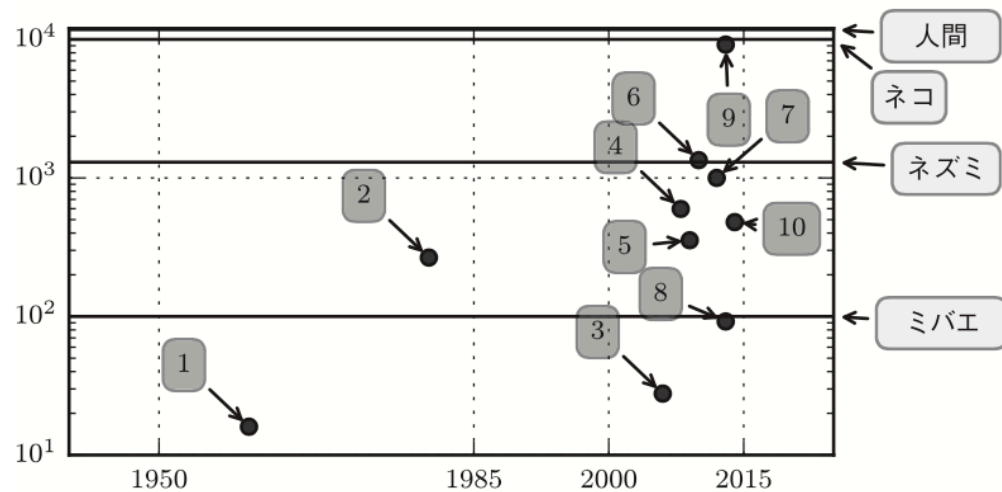
1.2.2 データサイズの増大

- デジタル化・インターネット化が進むにつれデータが増加
 - ベンチマークも、初期は手動で作られた数千のデータであったが、1980年には数万、2010年には数千万と大きくなってきた
- 2016年時点での必要なデータサイズの目安
 - カテゴリごとに約 5000 のラベルありデータ → 満足する性能
 - 少なくとも 1,000 万のラベルあり事例を含むデータ → 人間に匹敵

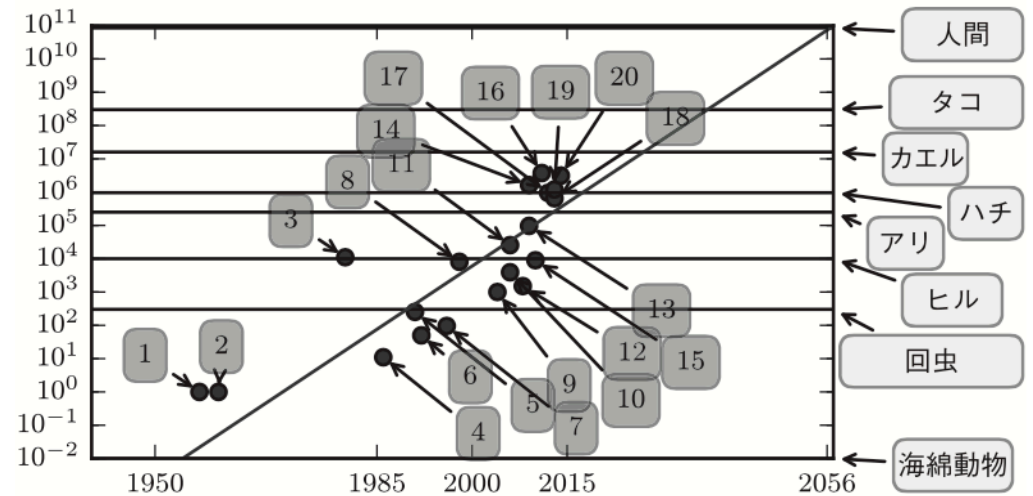


1.2.3 増大するモデルサイズ

- 計算機資源が豊富になりモデルを大きくできるようになった
 - ニューロンあたりの接続数：人間並み
 - ニューロン数：カエルにも及ばない
- 2.4年毎に2倍のペースで成長し、2050年代に人間の脳と同数に



ニューロンあたりの接続数



ニューロンの総数

1.2.4 精度・複雑生・実世界の影響の増大

- 物体認識
 - 画像の前処理：物体の周りでのクリッピング → 必要なし
 - クラス数：2 → 1000
 - データサイズ：白黒 → 高解像度のフルカラー
 - 精度：26.1% → 3.6%（道路標識の分類は人間超え）
- 音声認識：2000年以降向上なし → エラー率を半分に
- その他応用例
 - 文字認識
 - 機械翻訳
 - ニューラルチューリングマシン
 - 強化学習：Atariで人間並に
 - 新薬開発・脳の3Dマップを作る顕微鏡・原子の解析 等

参考文献

- Deep Learning
 - Ian Goodfellow, Yoshua Bengio, Aaron Courville
 - 日本語版
 - <https://www.amazon.co.jp/%E6%B7%B1%E5%B1%A4%E5%AD%A6%E7%BF%92-lan-Goodfellow/dp/4048930621>