IBM **Research**

米国から見た機械学習の今

IBM ワトソン研究所 井手剛
Tsuyoshi Idé, T.J. Watson Research Center, IBM Research

内容

- ある人工知能エンジンの挫折
- 深層学習の現在
- AI脅威論の現在
- AIと日本のこれから

AIの産業応用に関する楽観論は、今転換点に差し掛かっている: ヘルスケア応用で何が起きたか

(削除)

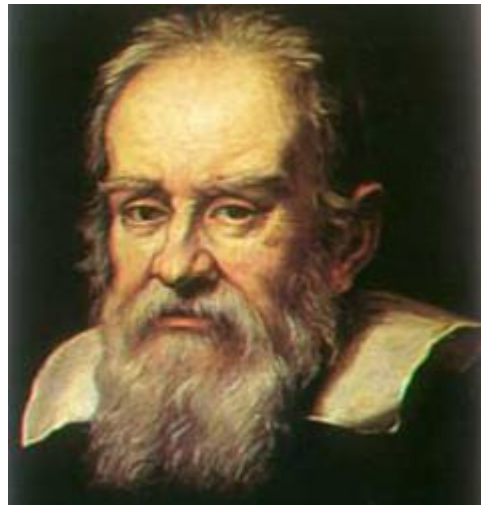
前回の「AI冬の時代」を振り返る:
古典的な人工知能研究の目標は、「エキスパートシステム」の構築

- エキスパートシステム ≒ 汎用人工知能
  - 任意の質問に、(専門家のごとく)答えてくれる機械
- エキスパートシステムの失敗は、以前のAI冬の時代の象徴
- 例: MYCIN (1970s)
  - スタンフォードで開発された医療診断システム
  - If-Thenルールの大きなデータベース
  - 論文上のベンチマークでは次々に好ましい結果が報告された
  - しかし実システムとして継続的に使われたものはほぼ皆無

セリーナ・ウィリアムスを起用したIBMの人工知能エンジン「IBM Watson」のテレビコマーシャル

前回の「AI冬の時代」を振り返る:「知識獲得のボトルネック」が最大の問題

- ほぼ完備なIf-thenルールの集合があれば、エキスパートシステムの技術的な課題は、ルールの高速検索技術に帰着できる
  - が、「ほぼ完備なIf-thenルールの集合」を作ることは現実的には無理なことが判明
- If-thenルールは、人間が理解するには便利な表現形式だが、「世界という本はIf-thenルールの言葉では書かれていない」
  - 天体の運動が、聖書の言葉では書かれていなかったのと同じ*
  - Galileo Galilei, "*The Assayer*", 1623.
    - ✓ “Philosophy is written in this grand book, the universe.... It is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures;....”
- 世界は、統計的機械学習の言葉で記述するのが（今のところはたぶん）自然



* T. Idé, Formalizing expert knowledge through machine learning. In S. K. Kwan, J. C. Spohrer, and Y. Sawatani, ed., Global Perspectives on Service Science: Japan, pp.157-175, Springer, 2016

今回何が違い、何が同じだったか

- 画像診断等の個別かつ特定のタスクで、深層学習が実用レベルに近づいたのは確かだが、一般的な診断問題を解くのはいまだに困難
  - 例: 過去の医学文献を網羅的に分析、与えられた条件から最適な診断を与える
  - AIに創造性は期待できないということ

- 深層学習がうまく働いている分野の特徴*
  - データ表現に曖昧性がない
  - 興味ある現象をほぼ網羅した学習データがある
  - タスクが広く知られており、多少間違ってもOK
- <u>この条件はほとんどの産業応用において当てはまらない</u>
  - データが現象を網羅しているかは普通はわからない。何を測定すれば網羅できるのかもわからないので、データ表現は常に曖昧
    - ✓ 頭痛が起きているのは、量子レベルの化学反応が原因かもしれないし、ミリ単位の血栓のためかもしれないし、肩こりのためかもしれないし、近くの風力発電所の低周波のせいかもしれない。
  - 医療診断や産業上の異常検知は、画像分類・翻訳・映画推薦 etc. と要求水準が全然違う。

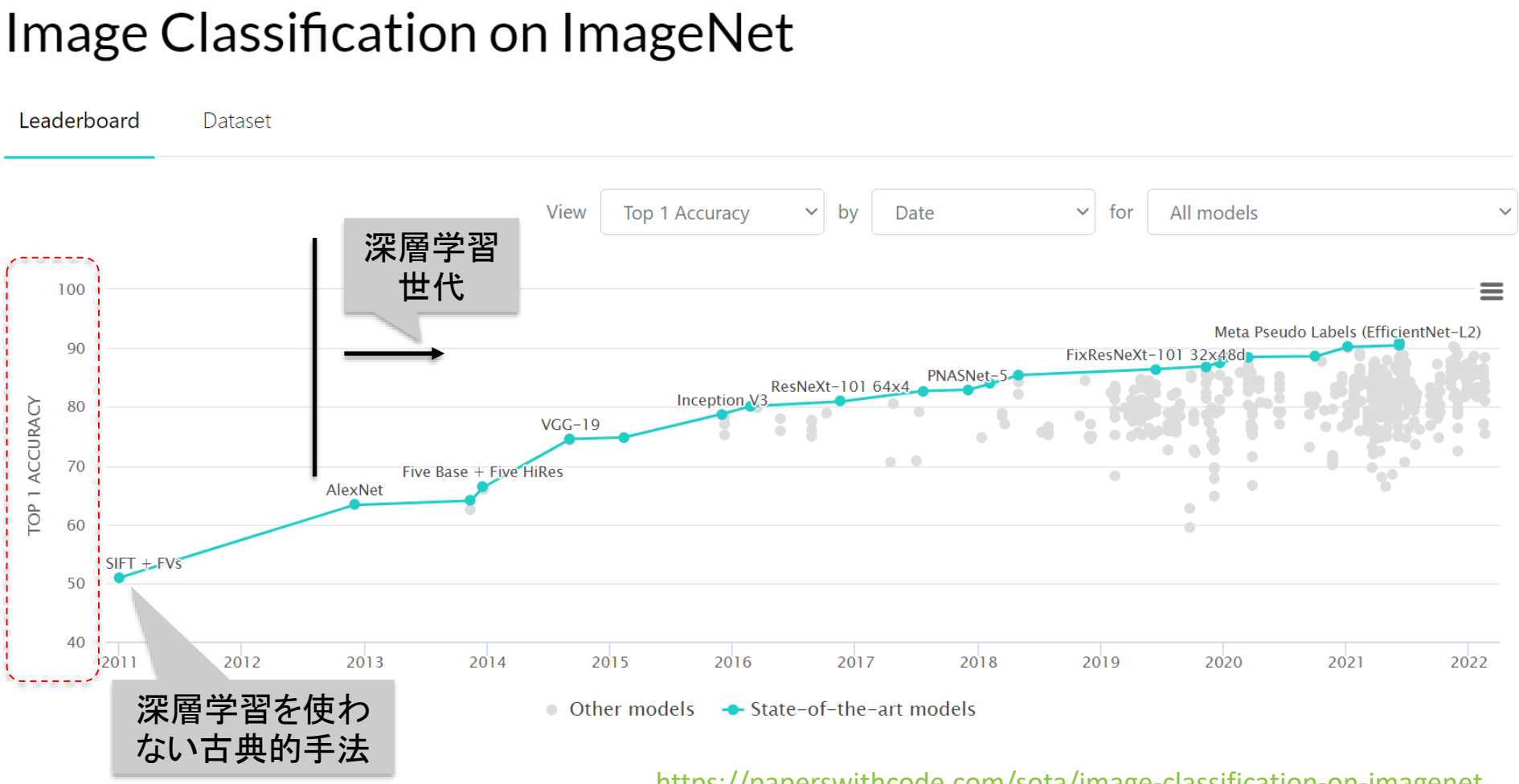
* T. Idé, “Recent advances in sensor data analytics,” Keynote talk at the 12th ICME International Conference on Complex Medical Engineering (CME 2018)

内容

- ある人工知能エンジンの挫折
- 深層学習の現在
- AI脅威論の現在
- AIと日本のこれから

画像認識の代表的なベンチマーク ImageNetの認識精度の推移
すでに人間の認識性能を凌駕

Top 1精度
（数千のカテゴリの中でランキングした時、トップに来たものが正解かどうか）



https://paperswithcode.com/sota/image-classification-on-imagenet

ImageNetベンチマークでの上位は「Transformer」と呼ばれる深層学習の方式が席巻している

- 自然言語処理、音声認識、その他のタスクでも同様の傾向
- Transformerとは何？
  - → 次ページ

Rank	Model	Top 1 Accuracy	Top 5 Accuracy	Number of params	Extra Training Data	Paper	Code	Result	Year	Tags
1	**CoAtNet-7**	90.88%		2440M	✓	CoAtNet: Marrying Convolution and Attention for All Data Sizes			2021	Conv+Transformer, JFT-3B
2	**ViT-G/14**	90.45%		1843M	✓	Scaling Vision Transformers			2021	Transformer, JFT-3B
3	**CoAtNet-6**	90.45%		1470M	✓	CoAtNet: Marrying Convolution and Attention for All Data Sizes			2021	Conv+Transformer, JFT-3B
4	**V-MoE-15B** (Every-2)	90.35%		14700M	✓	Scaling Vision with Sparse Mixture of Experts			2021	Transformer
5	**Meta Pseudo Labels** (EfficientNet-L2)	90.2%	98.8%	480M	✓	Meta Pseudo Labels			2021	EfficientNet, JFT-300M
6	**SwinV2-G**	90.17%			✓	Swin Transformer V2: Scaling Up Capacity and Resolution			2021	Transformer
7	**Florence-CoSwin-H**	90.05%	99.02%		✓	Florence: A New Foundation Model for Computer Vision			2021	Transformer

https://paperswithcode.com/sota/image-classification-on-imagenet

Transformer = 手作り特徴量入れまくりニューラルネットワーク

- 深層学習の初期の宣伝文句
  - 「ニューラルネットが自動で特徴量を見付けてくれる。人間を超える神業。特徴量工学はもう不要。全自動解析はもうすぐ」
- 2018年くらいから潮目が変わる
  - 系列予測問題（翻訳など）で、再帰ニューラルネットの（神秘的な）特徴抽出能力を忘れて、手作業的に特徴抽出部を組み込んだ方がよい
    - ✓ Vaswani, Ashish, et al. "Attention is all you need." Advances in neural information processing systems 30 (2017) →
  - ニューラルネットに昔のデータマイニング的工夫を組み込んだ感じ
    - ✓ 自己想起機構、単語位置情報の明示的組み込み、いろいろ混ぜて合議により予測

単語同士の自己相関を（RNNに頼らず）、明示的に計算する

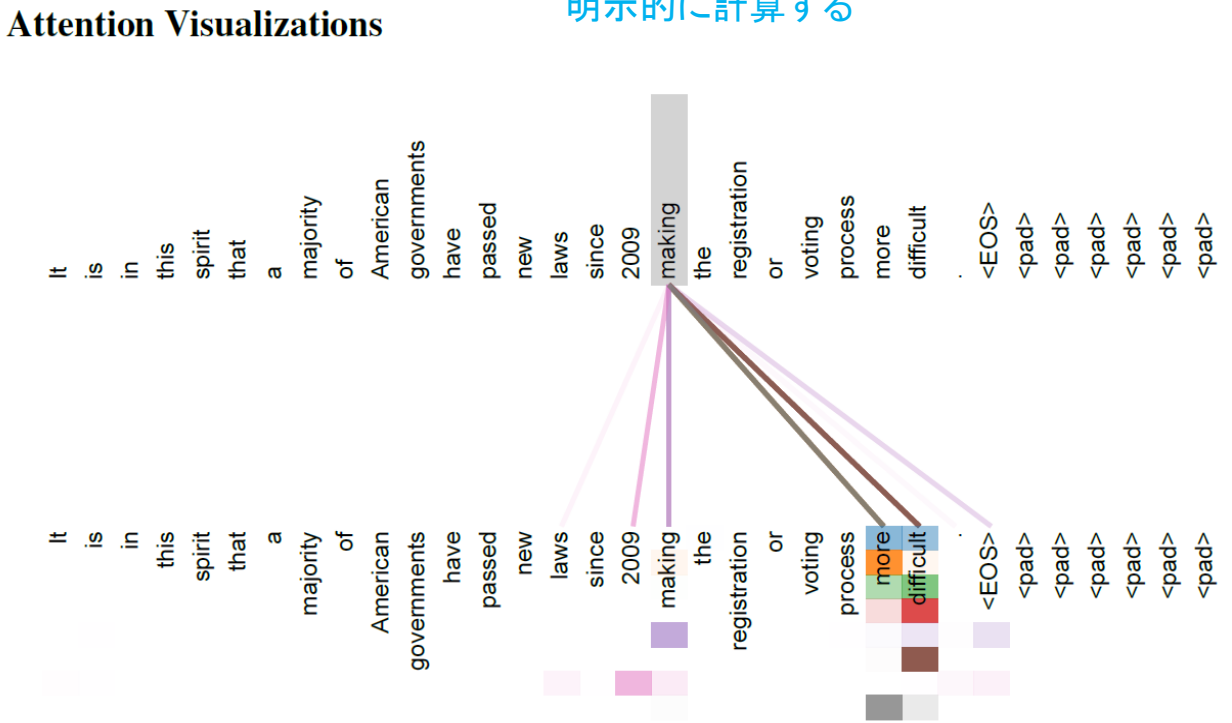


Figure 3: An example of the attention mechanism following long-distance dependencies in the encoder self-attention in layer 5 of 6. Many of the attention heads attend to a distant dependency of the verb 'making', completing the phrase 'making...more difficult'. Attentions here shown only for the word 'making'. Different colors represent different heads. Best viewed in color.

Transformerの（現時点での）成功が示唆すること

- 多くの実世界のタスクで、深層学習器（＝弱い非線形変換の多段階適用＋確率的勾配法）は、既存の汎用関数近似器より強力
  - 深層学習器 > フーリエ変換、直交関数展開、（再生核ヒルベルト空間での）カーネル展開、etc.
- 特徴量の工夫を併用するのはやはり有効
  - 一時期喧伝された全自動パターン認識とか「シンギュラリティ」とかは（現時点では）神話
- 深層学習は、何を入力として入れるか、という問いを解決できない
  - 画像認識、音声認識、自然言語処理では、長い研究の歴史の結果として「何をデータとして集めるか」という点に合意が確立している。深層学習のベンチマーク的成功はそれに乗っただけ。
  - ビッグデータは物理学の壁を絶対に越えられない
    - ✓ 例: iPhone で飛行機の写真を撮っても、亀裂進展による疲労破壊は予知できない
- 多くの実問題では、深層学習器の訓練・最適化・維持自体も大変

内容

- ある人工知能エンジンの挫折
- 深層学習の現在
- AI脅威論の現在
- AIと日本のこれから

「AIは社会的不正義を助長する」

- AIの科学研究への政治的介入は2022年時点ではますます強まる傾向
  - 機械学習のトップ会議 NeurIPS では論文の事前チェックが議論

- 論理的解決は困難
  - あらゆるデータは有限で、あらゆる変数の選択はある視点の表明である以上、いかなるモデルもバイアスを持つ

事前チェックをするとして、そのチェックリストの正しさはどう保証する？ 正しくないチェックリストにより生じうる被害はどう考える？歴史から何か学べる？

- ガリレオを弾圧したカトリック教会のチェックリスト
- ルイセンコ派を重用したスターリンのチェックリスト

「AIはブラックボックスであるがゆえ人間性の疎外を助長する」

- XAI (explainable AI) はAIの主要な話題のひとつ
- ふたつの大きな流派
  - 哲学派: 人間の主観的満足を追求
    - ✓ いまひとつ何がやりたいのかわからない
  - 工学派: AIの使い勝手を上げたい
    - ✓ まとも
- ヤン・ルカン教授の意見は傾聴に値する
  - 10章6節「AIは説明可能であるべきか」
    - ✓ 中身が分からないけど便利に使っているものは山ほどあるのに、なぜAIだけにより高い基準を設けるのか、という趣旨

産業応用の観点では、XAI研究は、アクション可能性の観点から行うのが健全

- 工学的には、説明とは actionable insights（人間の次の行動につながる情報）を適切に与えてあげること
  - 例: 奥さんの機嫌を推測するAI
    - ✓ 機嫌が分かればそれはそれで有用。
    - ✓ 加えて、「何をおれはやらかしたためかもしれないのか」が分かるとさらに有用
- しかし actionability は統計学者の興味の対象ではなかった
  - 例: 異常検知の古典手法としてのホテリングのT^2法
    - ✓ たとえば100次元のベクトルが与えられた時、「それ、全体として異常っぽいです」としか言ってくれない。x_2がまずかったのか、x_{74}だったのか、全然わからない
- AI脅威論の流れで出てくる XAI の議論の多くは主観的満足を優先しており、工学的問題解決よりもむしろ心理学実験に近い

内容

- ある人工知能エンジンの挫折
- 深層学習の現在
- AI脅威論の現在
- AIと日本のこれから

むすびに代えて

- 物理学、制御学と機械学習の協業は（いまだに）有望な研究領域
  - 画像・言語・音声での特殊事情は一般化できないことが多い。複雑系としての物理系の解析は相変わらず難しい（＝研究の価値がある）。
  - 汎用関数近似器としての深層学習をうまく使えば面白いことがまだできそう
    - ✓ 確定した問題領域での精度競争以外にもやることがたくさんある。
- AIの学会は中国の存在感が圧倒的。主要な技術の多くは中国人の著者による。トップダウンの戦略的投資がうまく回っているように見える
  - 政治・文化・社会の全分野において、ここ10年の日本の存在感の低下は（東海岸では）圧倒的
  - 日本社会および伝統的大企業の制度疲労は傍目には明らか
- 何にせよ、問題設定の作業を他人に任せないことが大切

おしまい

アクション可能性の観点からの最近の話題2つ

- ブラックボックス予測モデルの異常寄与度の計算法
  - “Anomaly Attribution with Likelihood Compensation,”
    - ✓ Tsuyoshi Idé, Amit Dhurandhar, Jiri Navratil, Moninder Singh, Naoki Abe,
    - ✓ In Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI 21, February 2-9, 2021, virtual), pp.4131-4138.
- 確率的イベントに対する因果分析
  - “Cardinality-Regularized Hawkes-Granger Model,”
    - ✓ Tsuyoshi Idé, Georgios Kollias, Dzung T. Phan, Naoki Abe,
    - ✓ Advances in Neural Information Processing Systems 34 (NeurIPS 21, Dec 6-14, 2021, virtual)

講演題目: 米国から見た機械学習の今
2022年3月9日（水）10:30-(JST), 20:30- (EST)

- https://mscs2022.sice-ctrl.jp/program#ss1
- 講演者
  - 井手剛、IBM ワトソン研究所
- 概要
  - それまで情報科学の地味な一分野に過ぎなかった機械学習は、過去数年の間に、他の学術分野のみならず社会一般にも何らかの影響を与えうる存在になった。言語・音声・画像における深層学習の成功は、人間のデータ分析スキルはもはや不要との見方を生み出し、マスメディアにおいてはAI脅威論がさかんに論じられた。やや視点を引いて眺めれば、これらの動きは、国際社会における米中の二極化と、その陰画としての日本の没落と軌を一にしているようにも見える。米国においてこれらの動きの中に身を置いてきた立場で、機械学習の今について個人的な見解を語ってみたい。