

AIの基本知識と今後の社会について

国立情報学研究所・総合研究大学院大学教授

やまだ せいじ
山田 誠二



1 AIの現状

2020年現在、人工知能 AI (Artificial Intelligence) は、2015年ごろから始まった第3次 AI ブームが少し落ち着き始めた時期にある。そして、これからますます企業や自治体での応用を初めとする社会導入が進んでいく段階にあると考えられる。

このことは、ガートナー・ジャパンのハイブ・サイクル (毎年 IT 関係の動きがどういうポジションにいるかを一周期の波にプロットしたグラフ図) の2019年版にもよく現れている。AI に対する社会の「過剰な期待」は2019年でピークを過ぎており、比較的なだらかに安定期 (そして、その先には「幻滅期」がある) に入ってきている。まさに、AI にとって、これからの4、5年が社会インフラ技術として生き残れるか否かの正念場にあると考えられる。我々研究者も身を引き締めて、この状況を深く認識することが重要である。

2 AIの簡単な歴史、3回のAIブーム

AIには種々様々な定義があるのだが、コンセンサスが得られていると筆者が考えるものは、『人間並みの知的な処理をコンピュータ上に実現したプログラム』という定義である。ここで重要なのは、AIは人格や自意識をもつ人間に代表される生物のような何かではなく、あくまでコンピュータのプログラムに過ぎないということである。

プログラムである AI は、その動きがいくら複雑になろうとも、人間によって書かれた命令を逐一その通りに実行しているに過ぎない。その意味では、ATM システムや HD レコーダーのシステムとやって

いることは変わらない。この認識が、特に AI を研究したことのない、あるいは開発したことのないエンドユーザにはとてもわかりにくく、また大きく誤解される点である。

換言すると、「人は AI を擬人化する傾向が強い」ということである。そのため、「AI が自分で進化する」、「AI に職を奪われる」というような、あたかも AI が人格をもった何かであるような表現や言説が一般に受け入れられてしまう。

AI の研究開発の歴史は古く、1956年のダートマス会議から始まっており、60年以上の歴史がある。毎年現れては消えていくバズワード的な IT 技術において、これほど長い歴史のあるものは、多くはないのではないと思われる。しかし、AI は古いが、いわゆる「枯れた」技術ではないのである。「枯れていない」技術とは、まだまだ課題がたくさんあり、十分に発展していない技術を意味する。AI が枯れていないことは、話題を少しでも広くすると、人間と真に対等に会話のできる AI がまだ存在しないことから、おわかりいただけるであろう。

図1に示すように、これまでの AI 60年の歴史において、ほぼ20年周期で3度の AI ブームがあった。最初の第1次 AI ブームは、1960年前後でダートマス会議の後に起こった。そして、1980年代には、第2次 AI ブームが起こった。この第2次 AI ブームでは、当時の通産省主導の AI 国家プロジェクト、新世代コンピュータ技術開発機構 ICOT が盛り上がったこともあり、日本が世界の AI 研究開発を牽引していたと言える。このときは、ルールベースで人間の専門家の推論をシミュレートするエキスパートシステムがたくさん開発された。そして、その後の15

年以上にわたる「AI冬の時代」を経て、ディープラーニングの台頭とGAFに代表されるアメリカのIT企業のAI応用が牽引する形で、2012年辺りから現在に至る第3次AIブームが起こったのである。

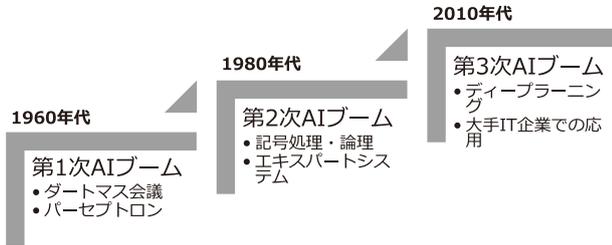


図1 過去3回のAIブーム

この第3次AIブームの特徴を一言で言うと「機械学習の実用化」が挙げられるだろう。計算機パワーの飛躍的な向上とハードウェアが安価になったこと、そして大量のデータであるビッグデータが利用可能になったことから、大規模な最適化や統計的な計算を必要とする現在の機械学習が十分な性能を発揮するようになったことが大きい。もちろん、様々なAIアルゴリズムや機械学習アルゴリズムの改良、そしてディープラーニングの発展も重要な牽引役となっていることは明らかである。

3 機械学習とは何なのか

少々技術的な内容になるが、機械学習 (machine learning) について触れる。機械学習とは、AIの一種であり、「学習できるAIを作る技術」と言っていだろう。機械学習を実現する方法にはたくさんの種類があるが、「教師あり分類学習」と呼ばれる分野が一番よく研究開発もされているので、それについて説明する。

教師あり分類学習とは、正解を教えてくる教師がいる状況 (実際には、正解付きデータ、つまり訓練データがある状況) で、その正解付きデータを入力して、AIが徐々にデータを2つのクラス (カテゴリー) A, Bに分類できるようになっていく機械学習を意味する。最初は入力データを分類できないが、

訓練データをどんどん入力すると、そこからクラスAのデータとクラスBのデータはどこが違うのかを区別する関数 (判別関数) を探索していく。そして、機械学習アルゴリズムは、その見つかった判別関数を出力する。

ディープラーニングとは別のアプローチをとる、典型的な機械学習アルゴリズムとして、SVM サポートベクターマシンがある。今、図2にあるような、簡単な分類問題を考えてみよう。2次元上にある黒い点と白い点で表したそれぞれの訓練データを左右に分けるような直線 (これが判別関数) を引いてくださいという問題である。人間にこの間に線を引いてくださいと問うと、ほとんどの人は、図にあるマージン (領域) の真ん中辺り (識別境界の辺り) に直線を引く。

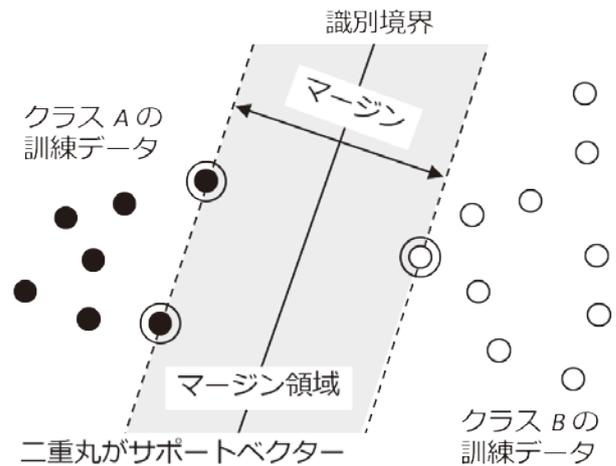


図2 SVM サポートベクターマシン

そこで、なぜあなたは真ん中に引くのですかと問うと、何かバランスがよさそうぐらいしか答えられないが、実はこの真ん中に引くというのがベストだということを数学的に証明したのが、サポートベクターマシン (の論文) である。

直線は、 $y=ax+b$ という方程式で表せる。ここで、 a は傾きで、 b は y 切片である。 a と b を決めればこの線が決まる。今度は先ほどの真ん中の線の a と b をどうやって計算するかが問題になる。SVMは、それらを非常に早く計算できる方法も与えている。

この線が引けると、次に新しいデータが出てきたときに、それが黒データなのか、白データなのかを自動的に判定できる。これは分類学習ができたことを意味する。

4 神経回路網とディープラーニング

機械学習のもう1つのアプローチは、人間が脳でやっている学習に学べというものである。それがニューラルネットワーク（神経回路網）と呼ばれる。今、流行のディープラーニングも、このニューラルネットワークの一種である。

人間の脳は数十億～100億個以上の神経細胞から構成されている。その神経細胞の間は、シナプス結合とよばれる結合方法で、導線のような結線が結ばれており、そこに電気信号が流れる。そして、その導線には、「重み」と呼ばれる可変抵抗のようなものがあり、その重みを変えることで、流れる電気信号が変化する。

実際の脳では、そのような結合が膨大にあり、ものすごく複雑な構造となっているが、数学的に扱うために、例えば図3のように簡単にモデル化する。

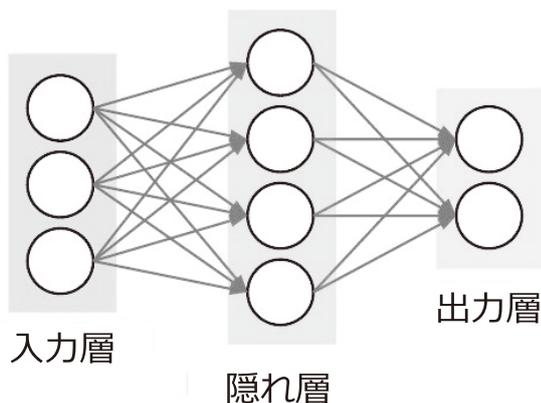


図3 ニューラルネットワーク

図3のモデルでは、入力層に信号が入力され、中間に隠れ層が1層あって、出力層がある。図中の○はニューロン（神経細胞）を、矢印はシナプス結合の重み付き結線を表す。今、犬と猫の画像を教師あり分類学習することを考えると、図3の出力層の2

つのニューロンのうち、上のニューロンは犬細胞、下は猫細胞というように割り当てられる。

電気信号は、入力層から出力層へ左から右に結線を伝って、重みをかけ算した強さで流れる。つまり、重みが大きいとたくさんの信号が流れる。そして、隠れ層のニューロンは、入力された電気信号を合計してそれがあるしきい値を超えると自分も右方向に結線を辿って電気信号を流すというメカニズムである。そして、出力層の犬細胞、猫細胞のどちらに多くの電気信号が流れてきたかで、犬か猫かを判定する。

ここで重要なのは、実際に電気信号を流すわけではなく、プログラムでシミュレートする点である。

学習はどのように行われるかという点、正しい入出力のペアである訓練データがたくさん与えられ、もし入力に対する出力結果が間違っていた場合は、修正するようにシナプス結合の重みを微調整していく。これが、ニューラルネットワークにおける学習であり、ディープラーニングも基本的には同様の学習方法を使っている。

図3では、隠れ層が一層のニューラルネットだが、この隠れ層はPCのメモリが許す限り、増やすことができる。そして、隠れ層が4、5層以上になると、ディープラーニングと呼ばれる。よく使われるのは、畳み込みニューラルネット CNN とプーリング層を組み合わせたものを幾重にも重ねていくもの(図4)であり、画像認識を中心に高い性能を示している。

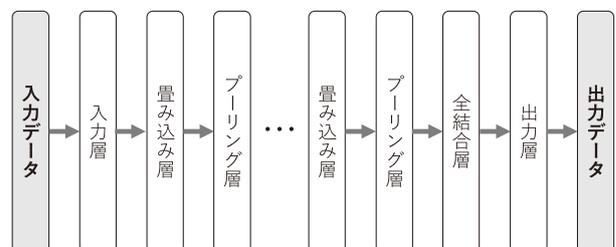


図4 ディープラーニング (CNN)

5 AIの得手／不得手

ここからは、AIの得意な分野と不得意な分野について説明する。図5に、それらをまとめている。

基本的には、変化の少ない（つまり、予測しやすい）世界で、外との情報のやり取りのない静的な世界は、AIが得意とする分野である。

換言すると、そのような得意分野には、AI導入の効果があるだろう。具体的には、事務作業のルーティンワーク、物理世界では屋内環境である。また、既に人間がAIに負けつつあるチェス、囲碁、将棋の類いのゲームである。

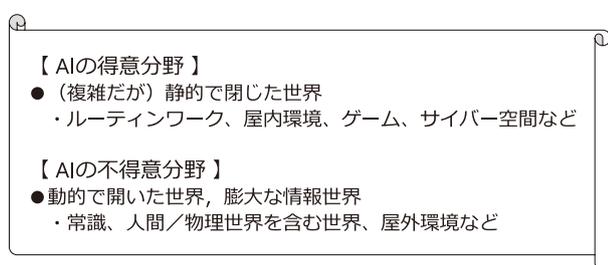


図5 AIの得手・不得手

これらの世界では、考える必要のある範囲（探索空間）が広くはあるが限定されており、AIでも効率的に調べることができる。また、コンピュータの方がはるかに広い探索空間を人間よりも高速に探索することができる。よって、AIの対象領域を十分に限定することで、AIが十分に能力を発揮できることになる。

例えば、自動車の自動運転は、一般道で任意の目的地という条件では、計算することが膨大になり、非常に難しいことが認識されつつある。しかし、例えば、短いルートを巡回するだけの路線バスであれば、静的で閉じた世界に近いので自動運転の可能性が飛躍的に上がると考えられる。

現在のAI、特に強化学習を始めとする機械学習は、AIが得意なテレビゲームの世界ばかりを対象に研究されている。正直ゲームが人間より強いAIができたところで、私たちの社会に与える貢献

はほとんどない。そこでは、AI研究者という専門家の悪い意味での現実離れが起こっている。このほかにも、常識の問題、人間を含む世界、簡単にだまされるAIなど、AIの苦手な世界についての議論はたくさん残っており、本来AIの目指している人間並の知能の実現と言うにはほど遠い。

6 進むAIの応用

AIの得手／不得手分野について議論したが、実はAIが応用できる課題はまだ膨大に残っており、それを開拓していくのは人間に他ならない。AI自身は、自分の解決できる課題を自分で見つけ出すことは不可能であり、それをやるのは、現場の人間やエンジニアなのである。

現在、特に医療の画像診断へのAI応用が進んでおり、注目されている。健康診断等で撮影する胃のX線画像から腫瘍部分を見つけるような画像診断である。この問題は、腫瘍の写っている画像と写っていない画像を分類できればいいので、前述の教師あり分類学習が適用できる。特に、CNNベースのディープラーニングは、この種の画像認識に高い性能を発揮するので、その効果が期待できる。実際、様々な応用研究がされている。

また、基本的には胃や大腸の内視鏡でのポリープ判定や顕微鏡画像での特定の細胞の発見も同じ問題なので、適用可能である。人間の代わりにAIが高速に大量の画像を見落としなく判定できるので、社会的なインパクトが非常に大きい。

一方、別の分野として、AIの導入が進んでいるのが、賢いAI情報検索を利用したヘルプデスク対応である。お客からの問い合わせを、最終的には全自動で行うことを目指している。ただし、現在やっているのは、質問をしてきた人の電話越しの音声認識して、過去に応答したことがあるパターンをAIが探す。そして、関連情報を即座にオペレータに表示して、オペレータがそれを見ながら答える。この方式の実現での開発が進んでいる。ここで、音

声認識と関連情報を上手く検索するところにAIが利用される。

また、一般事務処理では、現在RPA（Robotic Process Automation）の導入による自動化が進んでいるが、その先には、RPAのAI化が始まるだろう。人間のティーチングを認識する部分とAIプランニング技術によるRPA行動系列の完全自動生成が期待される。

7 AIによる労働観の変化、人間とAIの協働

AIと人間の労働について、1人の人間がやっている仕事を丸ごとAIが代替できるような言い方があるが、そんなことはあり得ないと考えられる。例えば、一見単純作業に見えるコンビニの店員さんの仕事の例を考えてみよう。実は、コンビニ店員の仕事は、20個以上のタスクに分かれ、その中にはAIではできないタスクがたくさんある。例えば、コンビニのヒット商品である「おでん」の仕込みは、AIにはとても難しい。おでんには、いろいろな固さの具がある。ゆで卵は堅いが、ゆで卵をつかむトングの強さで豆腐とか揚げをつかんだら潰れてしまう。そのことは、人間は見た瞬間に分かるが、AIがそれを計算するのは非常に難しい。それから、店やトイレや駐車場の清掃をすべて一台のAIロボットでやることも難しいと言える。

このように、1人の人間の仕事全部が丸ごとAIに置き換わるということは考えにくいですが、その一部のタスクは代替されるだろう。例えば、会社で企画書を作るときに、まず関連情報を集めるが、ネットから情報を集めて、それを取りまとめる作業はAIでもある程度は可能である。今、これらの前準備は人間がやっているが、その部分はもう人間がやらなくても済むようになる。そうなると、「本来人間がやるべき仕事はどこなのか」という意識が働くよう

になる。

従来、労働が、いつ、誰が、どこで、何をしたかというので分類されているとすると、次はAIで代替可能かどうかという新たな分類軸が導入されると考えられる。そして、この軸なしには、個人のやっている仕事を最適化できないし、組織全体の仕事も最適化できないだろう。

最後に、人間とAIの協働について述べる。今、チェスの世界で盛り上がっている対戦方法として、「アドバンストチェス」がある。この試合の方式では、人間とAIが混合チームになって、チーム対戦をするのである。つまり、人とAIが協働してチェスの試合をすることになる。また、驚くべきことに、現時点では、人間-AI混合チームの方が、最強AI単体チームよりも強いと聞いている。

さらに、このアドバンストチェスを考案したのが、IBMのAIに負けた元チェスのグランドマスターであったガルリ・カスパロフ本人であったことも、大変興味深い。このアドバンストチェスが、筆者の考える人とAIの協働を端的に象徴している。

このように、人間とAIが協働する場合には、「相互理解」が最も重要である。ここで言う「理解」とは、相手のモデルをもち、それをベースに相手の行動を予測できることを意味する。つまり、相互理解とは、お互いが相手のモデルをもって、相手の行動を予測できることである。

この人間とAIの相互理解を実現するには、AIは、人間が理解しやすいアルゴリズムで動くべきであり、かつ人間を理解するべきである。この2つの条件を満たすAIを実現することが、人間とAIの協働を成功させる重要なカギとなるだろう。この問題は難しく、まだ研究が緒に就いた段階であり、今後研究の進展が期待される。

参考文献

- ◎ 山田誠二（2019年）『本当は、ずっと愚かで、はるかに使える AI - 近未来 AI ロードマップ -』 日刊工業新社
- ◎ 山田誠二、小野哲雄（2019年）『マインドインタラクション』 近代科学社

寄稿者 PROFILE

山田 誠二（やまだ せいじ）

国立情報学研究所・総合研究大学院大学教授、東京工業大学特定教授

経 歴：1989年大阪大学大学院博士課程を修了後、同大学助手、講師、1996年東京工業大学助教授を経て、2002年より現職。専門は人工知能、HAI ヒューマンエージェントインタラクション。ここ10年の研究テーマは「人間と協調する人工知能」であり、現在 HAI、IIS 知的インタラクティブシステムを中心に様々な研究プロジェクトを推進中。人工知能学会前会長・顧問。

著 書：山田誠二（2019年）『本当は、ずっと愚かで、はるかに使える AI - 近未来 AI ロードマップ -』 日刊工業新社
山田誠二、小野哲雄（2019年）『マインドインタラクション』 近代科学社