

研究ノート

「AI・ロボット社会」における深層学習分析 （ディープ・ラーニング）・「自律型AI」が 引き起こす問題と法的対処の不備

浅野 美代子

目次

はじめに

1. CNNs：コンボリューションニューラルネットワーク
 - 1-1. 機械学習・人工知能：AI・ディープ・ラーニング
 - 1-2. CNNsによるデータ解析と問題点
 - 1-3. 実例：手書き文字解析における中間層可視化
 - 1-4. まとめ
2. 「AI・ロボット」とディープ・ラーニング
 - 2-1. AIの新しい発展：「自律型AI」と「AI・ロボット」
 - 2-2. AIの飛躍的な向上に対するディープ・ラーニングの役割
 - 2-3. まとめ
3. 「AI・ロボット社会」の現状と問題
 - 3-1. 今までになかった「AI・ロボット社会」
 - 3-2. 現状と今後の対応
 - 3-3. まとめ
4. 結論「AI・ロボット社会」への対応
 - 4-1. 「AI・ロボット社会」の取り組み
 - 4-2. 今後の対応

おわりに

はじめに

近年「AI」¹⁾という言葉が多くのところ語られている。それ以前には「ビッグデータ」²⁾が毎日のように語られていた。「AI」には、大量のデータが必要であるから含まれてしまったのか。すでに生活に入り込んでいる「AI」もある、例えば、自動翻訳、接客ロボット、自動掃除機などもある。これらは、一部のことをのみを行う「特化型AI」といえる。

「AI」によって新しい時代が到来するのであれば「AI」の正体を知って、一人一人が事前に対処するべきである。「特化型AI」、「汎用AI」、「自律型AI」などの言葉の意味も把握するべきである。

画像解析では、人間のレベルを越している所がある。この技術は自動運転で用いられ、すでに実用に供されている。画像分析には、Convolutional Neural Networks: CNNs³⁾モデル（深層学習：ディープ・ラーニング）が用いられているので、CNNsモデルを第1章で説明して、解析上での問題点もあげる。画像解析の他に会話、文章解析にもすでに利用されて、生活の向上になっている。

次に、AIとは何か、「AI」によって何ができて、何がかわるか、現状で生じる問題を論じる。

国内と海外での政策対応と自治体でのAI活用や実行、自主規制などの現状を述べて、政策と研究についても述べる。

今後の「AI」発展は予測できない点が多く、3年度、5年後、10年後をそれぞれで考えるのが妥当と思われる。

1. CNNs：コンボリューションニューラルネットワーク

LeCun, Y., Bengio, Y. and Hinton, G. E. (2015) によって、Deep Learningが提案された。畳み込みニューラルネットワーク：CNNのインスピレーションを得たのは、ネオコグニトロン⁴⁾である。これは、1980年代に福島邦彦

によって提案された階層的な多層ニューラルネットワークである。手書きの文字認識やパターン認識などに使用されている。福島の“neocognitron”は、1959年にHubelとWieselによる研究からヒントを得ている。ヒューベルとウィーセルは2つの業績により1981年ノーベル生理学・医学賞を受賞した。①1960年代から70年代にかけて行った視覚野に関する研究、②視覚からの信号がどのように大脳で処理され、形、動き、立体的な深さ、色などを検出しているかであり、視覚神経生理学を創始した。

CNNsモデルは、機械学習の最強の手法であり、多層ニューラルネットワークモデルで、中間層に畳み込み層と、プーリング層を含んでいる。1-1「機械学習・人工知能：AI・ディープ・ラーニング」で詳しくみる。「1-2「CNNsによるデータ解析と問題点」では問題点をあげて、1-3で「まとめ」を行う。

1-1 機械学習・人工知能：AI・ディープ・ラーニング

人工知能とは、人間が行う特定の作業あるいは活動全体について人間の模倣を目指して作られたものの総称で、一般に「Artificial Intelligence」を略してAIと呼ばれることもある。人間の活動全体を模倣しようとする「汎用型人工知能」と、人間の活動の一部を模倣する「特化型人工知能」の二つがある。「汎用型人工知能」は今のところなく、すべてが「特化型人工知能」である。

AI研究の発端は、今から60年ほど前の1950年頃で、イギリス人数学者アラン・チューリングが初めて人工知能の概念を提唱、その後アメリカ人計算機科学者ジョン・マッカーシーが公の場で「人工知能（Artificial Intelligence）」という言葉を用いた。

それから20年後、AIブーム再燃：人工知能に専門知識を取り込み、専門家と同様の判断が下せるようになった。機械学習に欠かせないものは、学習のために必要なデータであり近年、デジタル機器の急速な普及や通信技術の

発達で大量のデータ「ビッグデータ」が集められるようになった。機械学習では、データの特徴を与えるが、ディープ・ラーニングは、これまで人間が与えていたデータの特徴をAI自身が見つけ出す。

ニューラルネットワークは2012年にディープ・ラーニングがILSVRC (ImageNet Large Scale Visual Recognition Challenge)」と呼ばれる画像分類コンテストで、従来の機械学習の精度をはるかに上回る精度で優勝して以来、第3次ブームとなり全世界で注目されている。ディープ・ラーニングは画像認識・分類だけでなく音声認識や自然言語解析にも応用され、各分野でも目覚ましい研究成果が発表されている。

(1) ディープ・ラーニング

ディープ・ラーニングは、多層で構成された計算モデルであり、バックプロパゲーション・アルゴリズムを用い、学習によって内部パラメータを決定する。Convolutional Neural Networks: CNNs も含まれる。この方法は音声認識、視覚物体認識、物体検出、およびゲノム医療のような多くの他の領域における最先端技術を劇的に向上させた。

中間層の可視化で、各層の表現を前の層の表現から計算するために使用される内部パラメータをモデルがどのように変化させるべきかを示すので、データの複雑な構造を発見することもできる。

CNNsモデルの特徴

CNNは、畳み込みニューラル・ネットワークモデルであり、中間層が多層で、畳み込み層・プーリング層も含まれる。一番単純な3層フィードフォワード型は、中間層が1つの3層フィードフォワード型である。3層とは、入力層・中間層・出力層であり、隠れ層が1層のニューラル・ネットワークモデル (single-hidden-layer neural network) とも言う。図1参照。

入力層はデータで n 変数 $X = (x_1, x_2, \dots, x_p)$: 中間層の出力関数はシグモイド関数: $1 / \{1 + \exp(-x)\}$ である。 w_{ij} はウエイト: 学習できるパラメータ、したがって入力層のデータと出力は非線形関係になる。出力関数が線形

single-hidden-layer neural network

- Fit single-hidden-layer neural network
- $S(x)$ is sigmoid function such as logistic function, $1/\{1+\exp(-x)\}$.

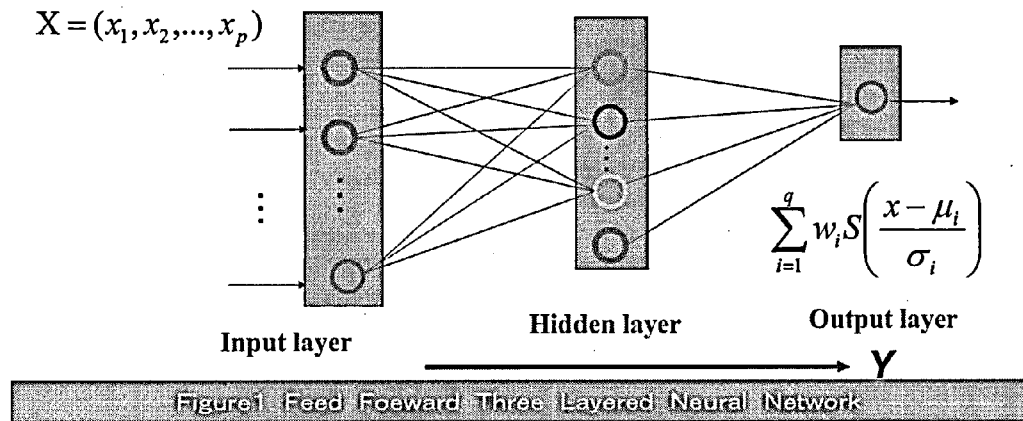


図 1. 3層フィードフォワード型ニューラル・ネットワークモデル

関数ならば、データと線形関係で重回帰分析となる。

3層フィードフォワード型ニューラル・ネットワークモデルは予測精度が線形回帰分析に比べて高い、しかしながら、中間層の役割がブラックボックスでモデルの解釈可能性が低いと言われている。これに対して浅野らはその原因を究明して、変化点への対処構造を発見した。その後ニューラル・ネットワークモデルと線形回帰分析のハイブリッドモデルを提案した。（Asano

The Hybrid model (Asano & Tsubaki, 2003) and the statistics

$$y = \beta_0 + \sum_{j=1}^p \alpha_j x_j + \sum_{i=1}^q \beta_i f_i \left(\sum_{j=1}^p w_{ij} x_j \right) + \varepsilon$$

$$= A + B + \varepsilon$$

- where

$$B = \sum_{i=1}^q \beta_i f_i \left(\sum_{j=1}^p w_{ij} x_j \right)$$

- $f(t) = \frac{1}{1+\exp(-t)}$ is the Sigmoid function
- under the condition that errors ε are independent, normal with constant variance

図 2. ニューラル・ネットワークモデルと線形回帰分析のハイブリッドモデル

& Tsubaki, 2003)⁵⁾ モデルを示す。

CNNsモデルは、機械学習の最強の手法であり、多層ニューラルネットワークモデルで、中間層に畳み込み層⁶⁾と、プーリング層⁷⁾を含んでいる。畳み込み層とプーリング層は、一对で使うことが良いという結果がある。「畳み込み層」は、画像からのエッジ抽出等の特徴抽出の役割を果たして、「プーリング層」は、抽出された特徴が、平行移動などでも影響を受けないようにロバスト性を与える役割を果たしている。次元も減らしている。CNNsモデル像解析例⁸⁾を示す。出力層では全結合分類⁹⁾である。

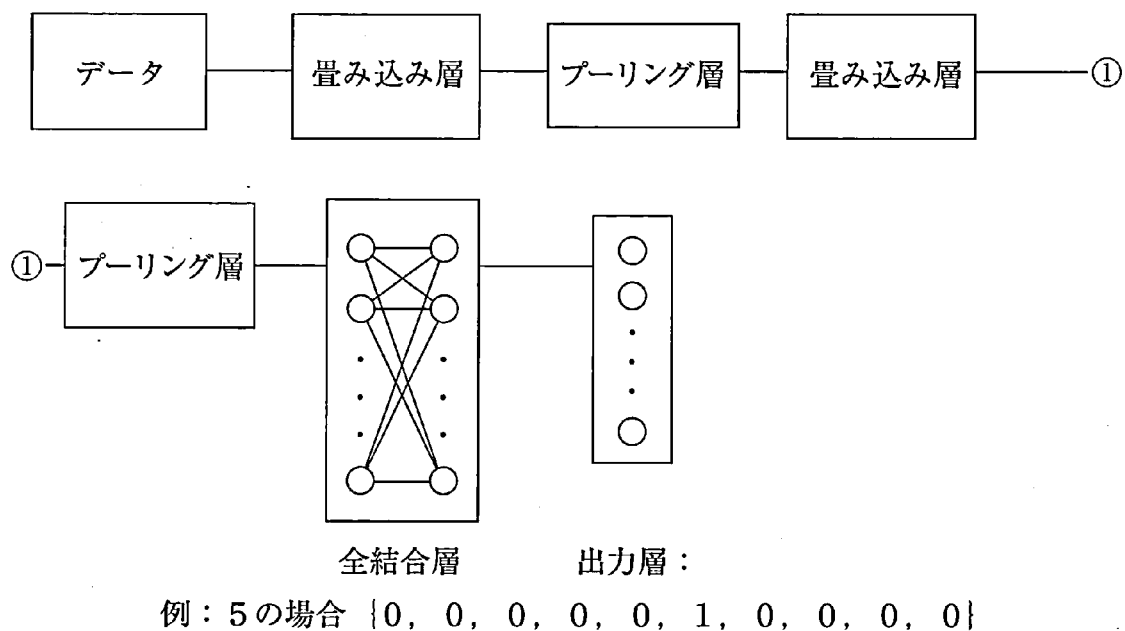


図3. 手書き文字解析コンボリューションニューラルネットワーク

1-2 CNNsによるデータ解析と問題点

(1) 入力データに落ちこぼれはないか？

入力データは大量のデータで分析を行う。いくら大量でもその中にないデータについては解析できない。すなわち構造が適正でない可能性がある。したがってウエイトも正しくない。実務に供されることは危険である。

(2) 作成したCNNsモデルは適正か？

データが完璧であっても求めた構造が最適でない場合がある。あらゆる検証用データを用意して実用に供される前に調べる必要がある。

(3) 収集されているデータの問題

現代社会においては、民間企業や行政機関により大量の個人データが収集、保有されている。また、大量の個人データがどのように分析され、誰によってどのように利用されているにつき、少なくとも当の本人は知っておくべきであろう。しかも、これらの大量の個人データの分析方法いかんによっては、個人データの「誤って」利用される危険性も否定できないものである¹⁰⁾。

しかし、自己のいかなる個人データが誰に収集され保有されているかを知る国民は少ない。この状況は、憲法が保障する基本的人権、特にプライバシー権との関係で重大な問題を孕むにもかかわらず、現状を見る限り、法や業界規制の場面でこの問題が十分に検討されていない。

1-3 実例：手書き文字解析における中間層可視化

Googleが2016年に公開した、誰でも無料で使える：オープンソース「TNSORFLOW」で、収録されている手書き文字データを用いてCNNモデルを実行する。手書き文字データ：訓練データと正解データ55000件、テストデータと正解データ5000件、検証用データ10000件 合計70000件から成る。データ構造は、データ： $28 \times 28 = 784$ ピクセル（0～255）、正解データ：0か1で10個、そのうちの1つが1。入力データの例を示す。

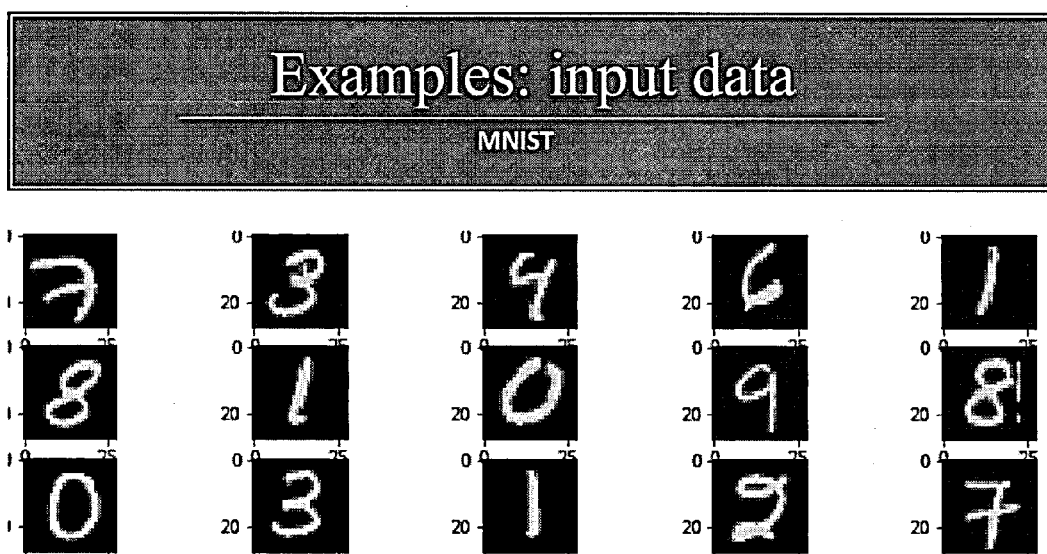


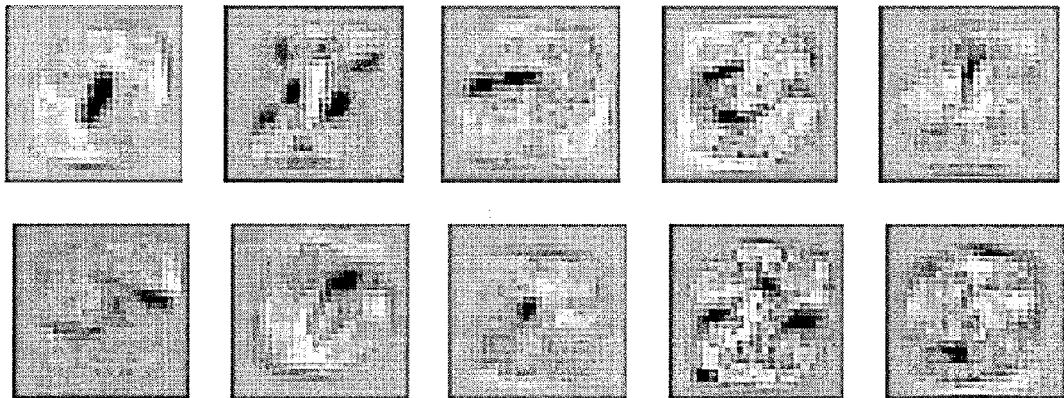
図4. Tensorflowによる手書き文字入力データ例

解析：CNNsモデル、入力層—中間層（畳み込み層・プーリング層・畳み込み層・プーリング層）—出力層

55000個のデータから TENSORFLOW の関数 `tf.nn.conv2d` を用いてウエイト w_{ij} を求める。ウエイト w_{ij} が分かれば、CNNsモデル内部での計算途中結果がわかる。中間層途中で何が起きているか知ることができる。

中間層の可視化例として、SOFTMAX回帰分析でのウエイトを示す。上段左から 0, 1, 2, 3, 4, 下段 5, 6, 7, 8, 9 のウエイトである。各々は $28 \times 28 = 784$ ピクセル (0~255)。なんとなく数字が書かれているように見える。手書き文字は 0 から 9 までの 10 種類の数値で結果も 10 種類と単純であるが、画像分析では、何が映っているのか、単体ではないので、中間層がもっと深く層の数が多くなる。画像が犬か猫かを予測する。自動運転では、他の車、信号機、壁、歩行者などを即座に判別している。

Tensorflow Visualization of the Soft-Max Regression: Weights 0—9



上段左から 0, 1, 2, 3, 4, 下段 5, 6, 7, 8, 9

図 5. 入力層から中間層へのウエイト w_{ij}

手書き文字データの SOFTMAX 回帰分析では、正答割合が 0.91 で、CNNs モデルでは、正答割合が 0.99 であった。

1-4 まとめ

ディープ・ラーニング（深層学習）は機械学習の最も強力な手法である。これは、畳み込みニューラル・ネットワークモデル（CNNs）である。すなわち、ニューラル・ネットワークモデルである。脳のデータ処理を模倣したモデルである。しかし中間層で何がおこっているかブラックボックスになってしまい、解釈可能性が低い。中間層の可視化によって構造を調べることができる。但し、中間層が20層になっている解析もあり、調べ尽くすことが困難である。

2. 「AI・ロボット」とディープ・ラーニング

2-1 AIの新しい発展：「自律型AI」と「AI・ロボット」

人工知能とは、人間が行う特定の作業あるいは活動全体について人間の模倣を目指して作られたものの総称で、一般に「Artificial Intelligence」を略してAIと呼ばれることもある。「自律型AI」は、外部から、リモートコントロールで操作されるのではなくて動くAIであり、「AI・ロボット」はロボットに「自律型AI」を加えたロボットを意味している。

2-2 AIの飛躍的な向上に対するディープ・ラーニングの役割

ディープ・ラーニング、すなわち、CNNsは予測精度が高い、これによって実務に供されることになった。例えば、画像による病気の発見では、人間のレベルと同等もしくは上になった。

2-3 まとめ

AIは予測精度が高くなって、実用に供された。そのためロボットに加えて「AI・ロボット」の可能性がでてきたことで、世の中に変化が生じている。

3. 「AI・ロボット社会」の現状と問題点

ディープ・ラーニングによって引き起こされたAIの飛躍的变化によって「AI・ロボット社会」が現実になる可能性が出てきた。AIはすでに生活の中に入ってきている。例えば、自動掃除機など。3-1 今までになかった「AI・ロボット社会」では、国内政策と海外の対応を述べる。3-2 現状と今後の対応について述べ3-3 でまとめる。

3-1 今までになかった「AI・ロボット社会」

いままでになかった社会が始まるとうしている。政策や海外政策ですすでに対策がとられている。

国内でのAI政策・自治体・民間対応を考察する。

(1) AI政策

内閣府はSociety 5.0 という考えを閣議決定した。これによって総務省・経済産業省・文部科学省は一緒に進むべき道を公開した。

A. 内閣府

総合科学技術・イノベーション会議のひとつとして「人間中心のAI社会原則検討会議」を開催している。

目的は、AIをより良い形で社会実装し共有するための基本原則となる人間中心のAI社会原則を策定し、同原則をG7及びOECD等の国際的な議論に供するため、産学官民による取組等を参考にしつつ取りまとめることである。

(産学官民の取組の一例)

1. 総務省AIネットワーク社会推進会議の「国際的な議論のためのAI開発ガイドライン案」

2. 人工知能学会の「倫理指針」

3. 日本経済団体連合会の「AI活用原則」

主な議論として、①AIが実装された社会、またはその前段階でのAI-Readyな社会を想定した「望ましい社会」、②各国が合意して皆で守ることができる原則を打ち出す会議のアウトプット、③AIで起こり得ることを想定した予防的なルールの整備、④技術を把握し、技術の未熟さも踏まえた原則の検討、等がある。¹¹⁾

B. 経済産業省

経済産業省は、平成30年6月、「AI・データの利用に関する契約ガイドライン」を策定した。AI技術とは、このガイドラインでは、主に「機械学習」を指している。

近い将来、AI技術が社会に広く普及し、利便性・生産性の向上や、今までは対応しきれなかった少子高齢化等の社会的・構造的課題への対応が可能になることが期待されている。しかし、AI技術を利用したソフトウェアの開発・利用に関して、権利関係をどのように処理すれば良いのか、開発・利用に伴って発生し得る責任を誰が負うのかといった法律問題は新しい問題であり、法律の整備が追いついていないのが現状である。

AI技術を利用したソフトウェア開発の依頼者（ユーザ）と開発者間の開発・利用契約を作成するにあたっての考慮要素、トラブルを予防する方法等についての基本的な考え方を提示。AI技術に係る契約の法整備が進むまでは、このガイドラインに沿った契約が主流になる見込みである。

また、新産業構造部会は平成28年4月、「新産業構造ビジョン」を発表した。現在、第4次産業革命の真ただ中である日本も重大な岐路に立たされている。就業構造が転換され、「仕事の内容」の変化が訪れる。

将来、「AIやロボットに代替され得る仕事」はいさぎよくAIに任せ、「AIやロボットを使って、共に働く仕事」に新たな雇用を生み出すチャンスであ

ると指摘している。経済産業省の試算では、製造・調達部門の従業員数は、約300万人減ると推測されている。一方の高級レストランなどの接客係など人間が行うことで高付加価値なサービスでは約180万人増えると予測されている。このように、AI・ロボットと共存していくための人間側の職業をシフトしていくことが重要になってくる。

C. 総務省

総務省情報通信政策研究所は、平成28年10月から「AIネットワーク社会推進会議」（議長：須藤 修 東京大学大学院情報学環教授・総合教育研究センター長）を開催し、AIネットワーク化をめぐる社会的・経済的・倫理的・法的な課題について検討を進め、同推進会議において、平成30年7月に『報告書2018 -AIの利活用の促進及びAIネットワーク化の健全な進展に向けて-』がまとめられ公表された。

「AI利活用原則案」の項目として、次の①から⑩が今後検討される。① 適正利用の原則、② 適正学習の原則、③ 連携の原則、④ 安全の原則、⑤ セキュリティの原則、⑥ プライバシーの原則、⑦ 尊厳・自律の原則、⑧ 公平性の原則、⑨ 透明性の原則、⑩ アカウンタビリティ¹²⁾の原則である。

D. 文部科学省

Society 5.0 の実現に向け、及びその進展の中で、広く国民にはどのような能力が必要か、また、社会を創造し先導するためにどのような人材が必要かについて、その社会像を具体的に描きながら議論するため、「Society 5.0 に向けた人材育成に係る大臣懇談会」を開催している。

「Society 5.0 に向けた人材育成～社会が変わる、学びが変わる～」平成30年6月5日、Society 5.0に向けた人材育成に係る大臣懇談会、新たな時代を豊かに生きる力の育成に関する省内タスクフォースとしての本文・概要・動画が公開された。以下の第1章、第2章と第3章で構成されている。

第1章 Society 5.0 の社会像と求められる人材像、学びの在り方

（「Society 5.0 に向けた人材育成に係る大臣懇談会」における議論を踏まえて）

第2章 新たな時代に向けて取り組むべき政策の方向性

（「新たな時代を豊かに生きる力の育成に関する省内タスクフォース」における議論の整理）

第3章 新たな時代に向けた学びの変革、取り組むべき施策

（Society 5.0 に向けたリーディング・プロジェクト¹³⁾）

（2）自治体による行動

自治体はこれをうけて協議会をつくり広報や教育を開始した。特徴は大学・研究所・民間企業の協働で研究する。

（3）その他（大学・研究所・民間との協働）

A. 国立大学法人京都大学・株式会社日立製作所¹⁴⁾

京都大学と日立は、2016年、日立未来課題探索共同研究部門（日立京大ラボ）を設立。こころの未来研究センター広井良典教授と3名の研究者は、文理融合共同研究の一環として、京都大学の人文・社会科学系研究部門の社会構想と政策課題に関する知見を株式会社日立製作の人工知能（AI）技術と融合させ、持続可能な日本の未来にはどのような政策が必要か提言を行った。解析は3部で構成されている。最初に「少子化」や「環境破壊」といった149個の社会要因についての因果関係モデルを構築、次に、AIを用いたシミュレーションにより2018年から2052年までの35年間で約2万通りの未来シナリオでシミュレーション予測を行い、これらを23種類の代表的なシナリオのグループに分類して、最後に各々のシナリオを分析を行った。

例えば、2052年に社会の①人口や出生率が低く、②財政や社会保障はよいが、③都市に人口が集中しているシナリオなど。これらは「都市集中型」と

「地方分散型」のシナリオで傾向が2分されたため、「都市集中型か、地方分散型か」、またその社会が「持続可能か、破局的か」の2つの観点で、シナリオどのように分岐するかという時期と要因解析を行った。持続可能な未来に向けて重要な社会要因とその時期を特定し、政策として提言された。この分析結果が公表されて多くの反響があった。

第2段階で日立のAIシステムを用いて可能なケースのシミュレーションが実行された。さらに、計算機能力も現在飛躍的向上しているという背景があり、結果の分類までもAIが行った。最終の分析は、AIでなく研究者が行った。従前の計算機環境ではケースを決めてシミュレーションを行って分析していた。新しい今回の分析は、大学と企業との協働で実現された。協働なくては実現できないことを考えると、協働は今後も大事な研究体制である。

2016年6月23日「京都大学と日立が「日立京大ラボ」を開設し、「ヒトと文化の理解に基づく基礎と学理の探究」を推進」より一部引用：

京都大学と日立は、日立京大ラボでの取り組みを通じて、「超スマート社会^{*1}」の実現（Society 5.0^{*2}）に貢献していきます。

*1 超スマート社会：必要なモノ・サービスを、必要な人に、必要な時に、必要なだけ提供し、社会の様々なニーズにきめ細やかに対応でき、あらゆる人が質の高いサービスを受けられ、年齢、性別、地域、言語といった様々な違いを乗り越え、生き活きと快適に暮らすことのできる社会。

*2 Society 5.0：サイバー空間とフィジカル空間（現実社会）が高度に融合した「超スマート社会」を未来の姿として共有し、その実現に向けた一連の取り組みのこと。狩猟社会、農耕社会、工業社会、情報社会に続くような新しい社会を生み出す変革を科学技術イノベーションが先導していくという意味を持つ。

B. 東京大学の取り組み¹⁵⁾

東京大学（総長：五神真）は、急速に発展し社会に大きな寄与と影響を与えつつある人工知能関連分野の先進的教育研究を強化すべく、総合的かつ体系的な取り組みを推進するため、2016年10月1日に次世代知能科学研究センター（Next Generation Artificial Intelligence Research Center）を設置して、2026年9月30日までの10年計画の研究教育活動を開始した。

現状の人工知能技術の枠組みとその限界を超え、真に人間のためになり、将来の社会、産業、経済、文化、学術を駆動する新たな次世代知能科学体系の構築と応用、それを踏まえた将来社会ビジョンの提示と実現、および先進的な教育体系の構築と先端人材育成を目指し、東京大学の多様な分野が文理を越えて融合し総合力を発揮する連携研究機構として活動している。

センターの目的を引用する。「現状の人工知能技術の枠組みとその限界を超え、真に人間のためになり、将来の社会、産業、経済、文化、学術を駆動する新たな次世代知能科学体系を整備し、その研究開発と社会実装を世界的に先導すること、人工知能関連技術の浸透を踏まえて目指すべき将来社会像を明らかにし、実現に向けて取り組むこと、および、当該分野の先進的な教育体系の構築と次世代を担う人材の育成を推進することであり、具体的には、「動の実世界知能：現在の深層学習、人工知能を超え、相互作用や動的情報を基本原理とした次世代知能」と「人間的人工知能：人間の心や価値観、倫理も理解する次世代知能」に焦点を当て、幅広く展開することを目指している。

AI、ビッグデータ活用、IoT、ロボット、等の人工知能関連技術が急速に発展し、社会にさまざまな効能と影響をもたらし始め、これに呼応する動きも産業界や公的機関等で加速している。しかし、将来の発展の方向性と社会のあり方に向けた総合的・体系的な取り組みは、未だ十分とは言えないので総合的かつ体系的な取り組みを推進。」と述べている。

また、東京大学生産技術研究所（研究代表者＝関本義秀准教授）は、情報

通信機構（NICT、東京都小金井市）からの研究委託を受け、2016年11月～2018年度末までの期間、「次世代型市民協働プラットフォーム“My City Report”」の研究開発を千葉市・室蘭市・足立区・墨田区・沼津市・弘前市で実施している。

株式会社WACUL（本社：東京都千代田区、代表取締役：大津裕史）は、人工知能の研究を行っている東京大学松尾研究室（研究室長：松尾豊准教授）と協働し、Web最適化の研究開発を行うことを発表（2015年8月20日）。

みずほ銀行が東大の研究グループらと共同で人工知能を活用した「外国為替取引高度化」の研究を開始すると発表（2017年7月4日）。共同研究は、IGPI ビジネスアナリティクス&インテリジェンスや東京大学大学院工学部技術経営戦略学専攻・松尾豊研究室と協働研究として実施。

その他にも、協働研究が数多く企画されて実行されている。

C. 慶応義塾大学の取り組み¹⁶⁾

AI社会共創・ラボ（AI & Society Laboratory）、2016年7月6日開設、代表者：新保史生総合政策学部教授、他3名。

科学技術振興機構（JST）社会技術研究開発センター（RISTEX）から発信する公募型研究開発領域「人と情報のエコシステム（HITE）」に「法・経済・経営とAI・ロボット技術の対話による将来の社会制度の共創」プロジェクトとして採択。HITEではビッグデータを活用した人工知能、IoT、ロボットなどの情報技術を、人間を中心とした視点で捉えなおすこと、そして一般社会への理解を深めながら、技術や制度を協調的に設計していくことを目指している、HITEへの参加を通し、AI・ロボット技術分野、社会制度に関わる方、倫理、経済、経営分野の研究者そしてAI・ロボットの利活用を先導する国内企業など、多様なステークホルダーとの対話を通し、未来の技術と社会のあり方を共創することを目的としている。加えて最先端のルール、社会基盤の構築に必要な制度を提案する計画である。

目的を引用する、「汎用性の高いAIの進化は近年めざましく、今後も更なる発展が見込まれる。急速な技術発展が社会へ与える影響を予見し、社会制度の議論を行うには分野横断的で多様な観点からの評価軸の設定が望まれる。

しかし、日本国内にはAI・ロボット技術が社会に与える影響を体系的に研究した試みはなく、技術発展を見込んだ新しい社会制度作りの準備が十分になされていない。本研究では、技術分野と社会制度に関わる法、倫理、経済、経営、さらに芸術・建築などの創造性に関わる分野の研究者、そしてAI・ロボットの利活用を先導する国内企業など、多様なステークホルダーとの対話を通し、未来の技術と社会のあり方を共創していくことを目的とする。文理融合の研究グループによる社会と技術の共進化促進のための研究を実現することにより、技術と人の共存を前提とした制度設計を行う基礎理論を構築することが可能となり、本研究の実施により将来の制度設計やルールづくりの方向性を示す。加えて世界の拠点として、最先端のルールや社会基盤の構築に必要な制度を検討及び提言を目指す。」としている。

その他に、協働研究が数多く企画されて実行されている。

D. 情報ネットワーク法学会¹⁷⁾

「情報ネットワークをめぐる法的問題の調査及び研究を通じて、情報ネットワーク社会の健全な発展に貢献すること」を目的として2002年7月27日に設立された法律系の学会に、研究会「ロボット法研究会」が設立された(2016年5月31日)。

E. 人工知能学会¹⁸⁾

人工知能学会倫理委員会¹⁹⁾(委員長 松尾豊、東京大学)では、「人工知能学会 倫理指針」を作成し、2017年2月28日の人工知能学会理事会での承認を経て公開を行った。

人工知能学会倫理委員会ホームページから引用：

2014年の委員会設置以来、人工知能研究あるいは人工知能技術と社会との関わりを広く捉え、それを議論し考察し、社会に適切に発信していくことを進めてまいりました。我が国でも、さまざまな政府機関で人工知能と社会に関する議論が行われ、また国際的にもそうした議論が進められるなか、人工知能学会としても深い専門知識に基づいて、国内の議論をリードしていく役割があると考えています。

これは、人工知能研究者が自ら研究しているものに対して、社会に対して大きな影響を与える可能性がある、という考えで作成した指針・宣言である。その内容を簡略化すると、以下1から9になる。(9以外の主語は、人工知能学会会員)

1. 人類に貢献して、安全への脅威は排除するよう努める。
2. 法規制を遵守し、他者に危害などを加えるような人工知能を利用しない。
3. 他者のプライバシーを尊重する。
4. 公平・平等に人工知能を利用できるよう努める。
5. 人工知能の安全性とその制御の責任を認識する。
6. 人工知能が社会に与える影響が大きいことを認識して、誠実に振る舞う。
7. 人工知能の持つ危険性を認識し、悪用されることを防止するよう努める。
8. 社会と常に対話し、学び、理解を深め、自己研鑽に努める。
9. 人工知能が社会の構成員となるためには、上記倫理指針を遵守しなければならない。

数年前までは、人工知能などは夢の技術であり、当分出来そうもなかったもので、このような倫理指針の必要性はなかった。上記9は、アシモフのロ

ロボット3原則のようなSFの世界でしかありえなかった。しかし、兵器で「敵」を殺す判断を人工知能に任せることなど、その判断基準さえ学習させれば、ロボット兵器は、技術的に現状でも可能である。

アイザックアシモフ（1976）『私はロボット』ロボット工学の三原則

- 一、ロボットは人間に危害を加えてはならない。また何も手を下さずに人間が危害を受けるのを黙視してはいてはならない。
- 二、ロボットは人間の命令に従わなくてはならない。ただし第一原則に反する命令はその限りではない。
- 三、ロボットは自らの存在を護らなくてはならない。ただしそれは第一、第二原則に違反しない場合に限る。

次に、海外でのAI政策・自主規制対応について考察する。

（1）²⁰⁾ 欧州

欧州科学・新技術倫理グループが欧州委員会に対し声明文を発表。九つの論理原則（①人間の尊厳、②自律性、③責任、④正義・公平・連帯、⑤民主主義、⑥法の詩碑とアカウントビリティ、⑦セキュリティ・安全性・心身の整合性、⑧データ保護とプライバシー、⑨サステイナビリティ）を提唱。2030年代に欧州で完全自動運転社会を実現していくための工程の一つとして、AI倫理ガイドラインを策定する。

（2）²¹⁾ 中国

2030年までに理論・技術・応用の全分野で世界の最高水準まで引き上げ、世界の主要なAIのイノベーションの拠点となることを目標に掲げた「次世代人工知能発展計画」を発表した。

中国電子技術標準化研究所は、2018年1月18日に国内外のAIに関する標準化の動向を取りまとめた白書を公表。内容は、技術面の進歩及びその標準

化と並行して、安全面、倫理面、プライバシーの政策、法、及び標準の策定が重要である。

(3) アメリカ²²⁾

ホワイトハウスは、2018年5月10日、米国がAIについて主導的立場を取るための必要な政策を議論するため、100名を超える政府機関、産業界、学術界の代表を集めたAIサミットを開催、内容は、国家レベルのAIのR&Dエコシステムの支援、AIの便益を享受するための人材育成、AIイノベーションに対する阻害要素の除去、特定領域におけるインパクトの高いAI応用の可能性について議論。また、研究開発の省庁間での優先順位を調整し、連邦政府のAI投資計画を推進するAI特別委員会設立の公表を行った。

(4) Facebook、Google、Amazon、Microsoft、IBM が AI を共同研究設立²³⁾

3-2 現状と今後の対応

画像解析におけるCNNモデルが実用に供されてから飛躍的にAIが向上し、計算機能力も近年飛躍的に上がってAIが導入された。

例えば、車の自動運転が可能になった。現状において、これらの新しい状況には法律は対応していない。したがって、「AI・ロボット社会」における「法」とは、契約法、不法行為法、刑法といった従来の法体系の中でどのように位置づけられるかを考察しなければならない。²⁴⁾

3-3 まとめ

日本では、2016年Society 5.0への移行が閣議決定され政策が検討されている。一方、民間でもAIを活用して実用化する協働の動きが起こっている。国民が置き去りにになっているが、教育政策も検討されている。しかし、Society 5.0 実行の為に、AIを用いた新しい教育に大きな変化があるならば、事

前に十分な検討がなされるべきである。海外では、新しい時代に対応している。

4. 結論「AI・ロボット社会」への対応

4-1 「AI・ロボット社会」の取り組み

コンピュータの飛躍的向上で、データを大量に持って、解析に用いることが可能になり、一方でCNNsにより画像データが解析できるようになった。解析と同時に構造も予め与えるのではなく学習して把握する。判別の精度が、人間よりも向上した。

一方、医学・薬学分野の研究によって、脳における記憶などのメカニズムが徐々に分かってきている。CNNsモデルも脳の研究から生まれた。長年の研究によって、ついに機械が「目」をもったといえる。これは有効な利用方法、例えば、医療で病気の診断に活用など多くあるが、ロボットに画像解析と自動的に判断するプログラムを乗せればロボット兵器となりうる。

4-2 今後の対応

CNNsモデルの中間層の可視化を見るだけでなく、モデルの構造を理解することが大事である。予測精度が良いというだけでモデルを判断するのではなく、解析された構造を調べて分析をすすめるべきであり、構造の分析研究を行うことが今後の課題である。

おわりに

長年の研究によって、ついに機械が「目」をもったといえる。今後もし「耳」にあたる研究がすすめば、機械が「耳」も持てることになる。意識や心はどうなるか、今の段階ではわからない。「自律型AI」は現在存在していない。

今後「AIロボット」に仕事をさせる事が省力化になるかもしれない。「口

ボット」と共存する時代、すなわち職場や家庭で「AIロボット」と共存する時代が来るかどうかはわからない。AIの発展に対し「ブレーキ」になってはいけないという意見と、「歯止めをかけなければならない」という意見がある。

Society 5.0 は、AI社会を推進している。しかし「AI・ロボット社会」は皆で考えるべきである。

最後に、浅野ゼミ：AI研究グループ：小岸琢巳さん、塩屋憲佑さん、山口寛之さん、杉澤大貴さんは、tensorflowの環境設定と、AIに関する議論に参加していただきました。ここに記して感謝申し上げます。

〔付記〕本稿は、2018年度 統計関連学会連合大会（2018年9月9日から2018年9月13日）中央大学後楽園キャンパスで講演報告を行った内容に、加筆して作成しました。浅野美代子「AI・ロボット社会」における深層学習分析（ディープ・ラーニング）・「自律型AI」が引き起こす問題と法的対処の不備 私たちは本当にAIを使いこなせるか？ 2018年度 統計関連学会連合大会講演報告集、p.338.

- 1) 人工的につくられた人間のような知識、ないしはそれをつくる技術
松尾豊『人工知能は人間を超えるか』p.45
- 2) ビッグデータについては確立した定義は存在しないが、ビッグデータは単に「大量のデータ」を意味するものでなく、今まで扱われてきたものとは異なるより大容量かつ多様なデータを扱う新たな仕組みを意味する。福岡真之介・桑田寛史・料屋恵美(2017)『IoT・AIの法律と戦略』、商事法務P.139
- 3) LeCun, Y., Bengio, Y. and Hinton, G. E. (2015). Deep Learning, Nature, Vol. 521, pp 436-444.
- 4) Fukushima, Neocognitron (1980). "A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position". *Biological Cybernetics* 36 (4): 93-202.
Fukushima, K. & Miyake, S. Neocognitron (1982). a new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition* 15, 455-

4693.

Fukushima, K.; Miyake, S.; Ito, T. (1983). "Neocognitron: a neural network model for a mechanism of visual pattern recognition". *IEEE Transactions on Systems, Man, and Cybernetics* SMC-13 (3) : 826-834.

- 5) Asano, M., Tsubaki, H., Yoshizawa, T. (2002). "Effectiveness of neural networks to regression with structural changes", *Applied Stochastic Models in Business and Industry*, Volume 18, Number 3, 189-195.
- 6) **Convolutional layers**, which apply a specified number of convolution filters to the image. For each sub region, the layer performs a set of mathematical operations to produce a single value in the output feature map. Convolutional layers then typically apply a ReLU activation function to the output to introduce nonlinearities into the model.
<https://www.tensorflow.org/tutorials/estimators/cnn>
- 7) **Pooling layers**, which downsample the image data extracted by the convolutional layers to reduce the dimensionality of the feature map in order to decrease processing time. A commonly used pooling algorithm is max pooling, which extracts sub regions of the feature map (e.g., 2x2-pixel tiles), keeps their maximum value, and discards all other values.
<https://www.tensorflow.org/tutorials/estimators/cnn>
- 8) <https://jp.mathworks.com/solutions/deep-learning/convolutional-neural-network.html>
より引用
- 9) **Dense (fully connected) layers**, which perform classification on the features extracted by the convolutional layers and down sampled by the pooling layers. In a dense layer, every node in the layer is connected to every node in the preceding layer.
<https://www.tensorflow.org/tutorials/estimators/cnn>
- 10) 山本龍彦 (2017) 『おそろしいビッグデータ 超類型化AI社会のリスク』, 朝日新聞出版
- 11) 人間中心のAI社会原則検討会議 運営要綱 :
<http://www8.cao.go.jp/cstp/tyousakai/humanai/1kai/siryol.pdf>
人間中心のAI社会原則検討会議活動状況 :
<http://www8.cao.go.jp/cstp/tyousakai/jinkochino/7kai/siryol.pdf>
- 12) AIネットワーク社会推進会議報告書 2018-AIの利活用の促進及びAIネットワーク化の健全な進展に向けて - :
http://www.soumu.go.jp/menu_news/s-news/01iicp01_02000072.html
- 13) 文部科学省 Society 5.0 に向けた人材育成に係る大臣懇談会 :
http://www.mext.go.jp/a_menu/society/index.htm
- 14) 株式会社日立製作所ニュースリリース (2017.9.5) :
<http://www.hitachi.co.jp/New/cnews/month/2017/09/0905.html>
2016年6月23日「京都大学と日立が「日立京大ラボ」を開設し、「ヒトと文化の理解に基づく基礎と学理の探究」を推進」
http://www.kyoto-u.ac.jp/ja/research/events_news/office/kenkyu-suishin/sankangaku-renkei/news/2016/documents/01.pdf
- 15) 東京大学松尾研究室と人工知能ベンチャーWACULが共同研究, Webサイト最適化に向けた人工知能の開発を目指す :

- <https://wacul.co.jp/pressrelease/matsuo-laboratory/>
みずほ銀行と東大の研究グループが「外国為替取引高度化」の共同研究へ（2017年7月4日）：<http://ai-biblio.com/articles/543/>
- 16) 慶応義塾大学SFC研究所：AI社会共創・ラボAI & Society Laboratory（2016年7月6日開設）：<https://www.kri.sfc.keio.ac.jp/ja/lab/sfc-ai.html>
- 17) 情報ネットワーク法学会：<http://www.in-law.jp/>
- 18) 人工知能学会：<https://www.ai-gakkai.or.jp/>
- 19) 人工知能学会倫理委員会：<http://ai-elsi.org/>
- 20) ドイツ
①<http://www.fujitsu.com/jp/group/fjm/mikata/column/fjm-topics2/021.html>
②<http://journal.jp.fujitsu.com/2018/01/12/09/>
- 21) 中国
①<http://www.fujitsu.com/jp/group/fri/report/newsletter/2018/no18-008.html>
②<https://nissenad-digitalhub.com/articles/ai-china/>
- 22) アメリカ
①<https://johoseiri.net/trumpAI/> ②<https://www.jetro.go.jp/biznews/feature/usiot.html>
- 23) Facebook、Google、Amazon、Microsoft、IBM が AI を共同研究（2016年9月30日）：<https://nakedsecurity.sophos.com/ja/2016/09/30/facebook-google-amazon-microsoft-and-ibm-team-up-on-ai/>
- 24) 弥永真生・宍戸常寿（2018）『ロボット・AIと法』有斐閣
ウゴ・パガロ（2018）『ロボット法』新保史生（監訳・訳）勁草書房。
平野晋（2017）『ロボット法—AIとヒトの共生にむけて』弘文堂
福岡真之介・桑田寛史・料屋恵美（2017）『IoT・AIの法律と戦略』商事法務。