



白書

産業間の人工知能

※本冊子は、一般財団法人日本規格協会が平成30年度経済産業省の委託を受けて作成しました。

概要

人工知能は現在、産業界、研究者、政府及び投資家から多大な関心と注目を集めており、彼らは新しい機械学習技術とアプリケーションの開発に記録的な金額を投じている。作業の予測だけでなく、社会、ビジネス及び個人に影響を及ぼす現実的な決定を行う際に、人間の活動を支援するため、ますます高度なアルゴリズムが採用されつつある。製造部門で、ロボットが人間と共に作業するため自らの挙動を適合させている場合であれ、家庭環境で、冷蔵庫が住宅所有者の好みに基づいて食品を注文する場合であれ、かつては人間の熟練、判断力又は意思決定に委ねられていた領域に、人工知能は継続的に入り込んでいる。

人工知能は、環境資源の枯渇、世界人口の増加と高齢化、又は貧困との闘いなど、最も急を要する人類の幾つかの課題に対処する助けになる可能性をもつ一方、人間が適切な決定を下す助けとなる機械の使用が増えることは、ビジネス、政府及び政策決定者が理解し、注意深く取り組む必要がある多くのリスクと脅威を生み出している。安全、セキュリティ、プライバシー、信頼、及び倫理的な考察一般に関する新たな懸念が、人工知能が可能とする技術革新と共に確実に出現している。これらの課題は世界の全ての社会に共通するものであり、国際レベルで対処する必要がある。

この白書は人工知能が現在どのような地点に立っているか、及び今後5~10年における開発の見通しはどのようなものかを理解する枠組みを提供する。白書は現在の技術的能力の説明に基づき、現在使用されている主なシステム、技術及びアルゴリズムを記述し、それらが一般にどのような種類の問題を解決する助けとなるかを示している。産業の視点から、白書は人工知能技術の展開の広範な機会を提供す

る四つのアプリケーション領域、すなわち、スマートホーム、インテリジェント製造、スマート輸送と自動運転車、及びエネルギー部門について、より詳細に考察している。

これら四つの領域に関連する様々な具体的使用事例の分析は、人工知能が広範な業界に適用でき、それらに恩恵をもたらし得ることの明確な証拠を与える。この可能性は、現在及び将来のメガトレンドにより生み出される課題を解決する際に、人工知能がその方程式の本質的な部分になる道を開いている。白書はこの分析に基づき、人工知能が対処しなければならない既存の、及び将来の主な課題の幾つかを詳しく記述している。産業及び研究コミュニティは、データ、アルゴリズム、ハードウェア及びコンピューティングインフラストラクチャに関する技術的課題に取り組むための、イニシアチブを進展させる主な推進要因であり、一方、政府と規制当局は、人工知能の副産物になることが予想される最も重要な倫理的及び社会的問題の幾つかに対処するため、新しい方針を早急に練り上げる必要がある。

標準化と適合性評価は、市場における人工知能の採用を促進するだけでなく、機械による意思決定に関連する最も急を要する問題の幾つかを軽減する上で、不可欠な役割を果たすと期待されている。IECは、産業及び情報技術システムに対し、標準化と適合性評価機能の独自のミックスを提供する先導的組織として、国際レベルでこれらの課題の幾つかに対処する理想的なポジションにある。

IEC とその委員会で対象とされている次の具体的な推奨事項は、白書の最後の部分に記載されている：

- 水平的な人工知能標準化において、JTC 1/SC 42 の中心的役割を促進する。
- データの意味論(セマンティクス)と存在論(オントロジー)の標準化を調整する。
- 人工知能関連の使用事例を発展させ、それらを集中管理する。
- 一貫性のあるインタフェースを備えた人工知能基準アーキテクチャを開発する。
- 人工知能の適合性評価が必要な可能性があるか探る。
-
- 人工知能に関して、様々な社会的ステークホルダーとの対話を増す。
- IEC を含むテストベッド(試験台)において、人工知能の使用事例を含める。

人工知能は多くの異なる産業に渡る中核的技術となり、来るべき第4次産業革命の推進力の一つなることが予想されており、標準化団体はその将来を形作る上で重要な役割を果たす。IEC は、安全性と信頼性におけるその長年の実績に基づいて、この目標を達成し、人類への恩恵としての人工知能の見込みを実現させることに役立ち得る。

.....
謝辞

この白書は、プロジェクトパートナーであるドイツ人工知能研究センター(DFKI)、及びそのプロジェクトリーダーであるハイアールグループの多大な貢献の下、IEC 市場戦略委員会(MSB)の人工知能プロジェクトチームにより作成された。

プロジェクトチームは、ハイアール家電産業グループの副社長であり、MSB メンバーである Mr. Ye Wang によって指揮された。プロジェクトチームは次に記載する通りである：

Mr Jens Popper, DFKI / SmartFactory^{KL}, プロジェクトパートナーリーダー

Mr Jesko Hermann, DFKI / SmartFactory^{KL}

Mr Kai Cui, Haier, プロジェクト管理者

Mr Simon Bergweiler, DFKI

Dr Stephan Weyer, DFKI

Prof. Dr-Ing. Martin Ruskowski, DFKI

Mr Miao Wang, Haier

Mr Liang Guang, Huawei

Mr Yun Chao Hu, Huawei

Dr Victor Kueh, Huawei

Mr Di Wang, Huawei

Ms Mary Carol Madigan, SAP

Dr Ian Oppermann, NSW Data Analytics

Mr Sung-Min Ryu, KEPCO

Mr Lagyoung Kim, LG Electronics

Dr Sha Wei, CESI

Mr Ruiqi Li, CESI

Mr Xiaohui Du, ITEI

Mr Yujia Shang, ITEI

Mr Xiangqian Ding, OUC

Mr Guangrui Zhang, OUC

Dr Gilles Thonet, IEC

.....

目次

概要	3
略語リスト	9
用語集	11
第1節 序文	13
1.1 人工知能：奇跡か，屋気楼か？	13
1.2 人工知能の冬の時代から再生まで	14
1.3 大きな機会はリスクと課題を伴う	15
1.4 人工知能の定義	16
1.5 この白書の適用範囲	17
1.6 この白書の概要	18
第2節 人工知能のニーズ	21
2.1 天然資源の欠乏	21
2.2 気候変動	22
2.3 人口動態の傾向	22
2.4 経済政策	24
2.5 サービスと製品のカスタマイズ	24
第3節 人工知能を可能の実現要因と推進要因	27
3.1 人工知能の実現要因	27
3.1.1 計算能力の向上	28
3.1.2 データの利用可能性	29
3.1.3 改良されたアルゴリズム	29
3.2 人工知能の推進要因	30
3.2.1 クラウド及びエッジコンピューティング	30
3.2.2 モノのインターネット	31
3.2.3 ビッグデータ	31
3.2.4 消費者の受容	32

第4節	人工知能の内部	35
4.1	機械学習のカテゴリ	35
4.1.1	教師あり学習	35
4.1.2	教師なし学習	36
4.1.3	強化学習	36
4.2	現在の機械学習システム	36
4.2.1	コンピュータビジョン	36
4.2.2	異常検出	37
4.2.3	時系列分析	37
4.2.4	自然言語処理	37
4.2.5	推薦者システム	38
4.3	機械学習のためのアルゴリズム	38
4.3.1	決定木	38
4.3.2	サポートベクターマシン	38
4.3.3	ナイーブベイズ	40
4.3.4	k近傍法	40
4.3.5	k平均法	41
4.3.6	隠れマルコフモデル	42
4.3.7	人工ニューラルネットワーク	42
4.3.8	畳み込みニューラルネットワーク	43
4.3.9	リカレントニューラルネットワーク	43
第5節	人工知能の配備	45
5.1	スマートホームにおける人工知能	46
5.1.1	スマートテレビ制御システム	46
5.1.2	浴室のセルフサービスシステム	47
5.1.3	インテリジェント食品識別システム	47
5.1.4	スマートホームの課題	48
5.2	スマート製造における人工知能	48
5.2.1	予測保守	49
5.2.2	協力ロボット	51
5.2.3	品質管理	51
5.2.4	スマート製造における課題	53

5.3	スマート輸送と自動車部門における人工知能	53
5.3.1	自律運転	53
5.3.2	交通管理	54
5.3.3	交通ロボット	56
5.3.4	スマート輸送における課題	56
5.4	スマートエネルギーにおける人工知能	56
5.4.1	グリッドの管理と運用	57
5.4.2	消費者の関与とサービス	57
5.4.3	統合スマートエネルギープラットフォーム	57
5.4.4	スマートエネルギーにおける課題	58
第6節	人工知能の課題	61
6.1	社会的及び経済的課題	61
6.1.1	意思決定の変化	61
6.1.2	先進的なサプライチェーンの運用	61
6.2	データ関連の課題	62
6.2.1	訓練データの選択	62
6.2.2	標準化データ	63
6.3	アルゴリズム関連の課題	64
6.3.1	ロバストネス	64
6.3.2	転移学習	65
6.3.3	解釈可能性	65
6.3.4	目的関数	66
6.4	インフラストラクチャ関連の課題	66
6.4.1	ハードウェアのボトルネック	66
6.4.2	プラットフォームとフレームワークの欠如	67
6.5	信頼性関連の課題	67
6.5.1	信頼	67
6.5.2	プライバシー	68
6.5.3	セキュリティ	68
6.6	規制関連の課題	68
6.6.1	責任	68
6.6.2	プライバシー	69
6.6.3	倫理	69

第7節	人工知能における標準化ギャップ	71
7.1	人工知能における標準化の取り組み	71
7.1.1	ISO/IEC JTC 1	72
7.1.2	IEC	73
7.1.3	ISO	73
7.1.4	ITU	73
7.1.5	IEEE	74
7.1.6	ETSI	74
7.1.7	中国における標準化作業	75
7.1.8	米国における標準化作業	75
7.1.9	欧州 AI アライアンス	76
7.1.10	コンソーシアム及びその他の組織	76
7.2	標準化のギャップ	77
7.2.1	調和的なデータモデルとセマンティクス	77
7.2.2	データモデルに基づく一般的なオントロジー	77
7.2.3	人工知能アルゴリズムの検証	77
7.2.4	人工知能のインフラストラクチャのベンチマークと評価	78
第8節	結論及び推奨事項	79
8.1	産業界に対する推奨事項	79
8.2	規制に関する推奨事項	80
8.3	IEC とその委員会に対する推奨事項	81
附属書 A	将来の発展	83
A.1	生物学に触発された人工知能	83
A.2	人間/人工知能の相互作用	83
A.3	人工知能利用デジタルツイン	84
A.4	自動化された機械学習	84
参考文献 (邦訳では省略)		

略語リスト

技術及び科学用語

AI	人工知能
AIR	自動画像認識
AMI	先進的な計量インフラストラクチャ
ANN	人工ニューラルネットワーク
API	アプリケーションプログラミングインタフェース
ASIC	特定用途向け集積回路
CNN	畳み込みニューラルネットワーク
CART	分類及び回帰ツリー
CPU	中央処理装置
DNN	ディープニューラルネットワーク
EBL	説明に基づく学習
FPGA	フィールドプログラマブルゲートアレイ
GDPR	(EU)一般データ保護規制
GPU	グラフィックス処理装置
GRU	ゲーテッドリカレントユニット
HMM	隠れマルコフモデル
HTM	階層型時間メモリ
ICT	情報通信技術
ID3	反復ダイコトマイザー3
IoT	モノのインターネット
IT	情報技術
k-NN	k近傍法
KPI	主要業績評価指標
LSTM	長期短期記憶
NLP	自然言語処理
NPU	ニューラル処理装置
RDBMS	リレーショナルデータベース管理システム

ReLU	整流リニアユニット
RNN	再帰型ニューラルネットワーク
SME	中小企業
SVM	サポートベクターマシン
TPU	テンソル処理装置

.....

組織，機関及び企業

CESI	中国電子技術標準化研究院
DFKI	ドイツ人工知能研究センター
EC	欧州委員会
ENI	(ETSI ISG)体験型ネットワークインテリジェンス
ETSI	欧州電気通信標準化機構
FG-ML5G	(ITU-T)5G を含む将来のネットワークの機械学習に関するフォーカスグループ
IDC	国際データコーポレーション
IEEE	電気電子技術者協会
ISG	(ETSI)業界仕様策定グループ
ISO	国際標準化機構
ITEI	機械工業儀器儀表綜合技術經濟研究所(中国)
ITU	国際電気通信連合
ITU-T	(ITU)電気通信標準化部門
JTC	合同専門委員会
KEPCO	韓国電力株式会社
NHTSA	国家道路交通安全局(米国)
OUC	中国海洋大学
SAC	中国国家標準化管理委員会
UN	国連

用語集

アプリケーションプログラミングインタフェース

API

様々なソフトウェア構成要素の間の明確に定義された通信方法から構成されるインタフェース

特定用途向け集積回路

ASIC

特定の目的のために、特定の一連の動作を行うことに特化した電子回路

注記 アプリケーション分野は、そのアーキテクチャにより定義されているので変更できない。

人工知能

AI

問題解決、学習、パターン認識を含む知的挙動をコンピュータで模擬する、コンピュータ科学の一分野

人工ニューラルネットワーク

ANN

実行が明示的にプログラムされることなく、訓練事例を与えることによってタスクを実行する目的で、コンピュータ科学でしばしば使用される、生物のニューラルネットワークに触発されて生まれた数学的構造

中央処理装置

CPU

コンピュータプログラムの命令を実行する電子回路

畳み込みニューラルネットワーク

CNN

通常、画像認識などのタスクに適用される特別なフィードフォワードネットワーク

ディープラーニング

ディープニューラルネットワークを使用する機械学習の分野

ディープニューラルネットワーク

DNN

入力を処理して出力するために接続される、ニューロンの幾つかの連続層をもった人工ニューラルネットワーク

説明に基づく学習

EBL

訓練事例から一般化を行うためにドメイン理論を使用する、人工知能の一形態

フィールドプログラマブルゲートアレイ

FPGA

様々なアプリケーションに特化して動作する電子回路

注記 特定用途向け集積回路とは対照的に、FPGAは製造後に再プログラミングすることができる。

一般データ保護規制

GDPR

欧州連合における、データ保護とプライバシーに関する一連の重要な規制の変更であり、人工知能システムによる自動意思決定も扱っている。

グラフィックス装置

GPU

大量の計算を並列的に実行して画像を処理することに特化した電気回路

隠れマルコフモデル

HMM

マルコフ過程を使用して記述することが可能な線形シーケンスの確率モデル

注記 隠れマルコフモデルは、記述されたプロセスの全ての状態は直接観察できずそれゆえ隠されているという仮定の下、機械学習で使用する手法である。

モノのインターネット

IoT

物理デバイスのネットワーク，組み込みの電子機器，あるいは，データ交換のためにこれらのコンポーネントをより大きなネットワークに接続できるようにするソフトウェア

機械学習

実行を明示的にプログラムすることなく，データを増やすことによりデバイスが特定のタスクのパフォーマンスを改善できるようにする，コンピュータ科学のアルゴリズムのカテゴリ

自然言語処理

NLP

音声認識，言語理解又は言語生成のため，コンピュータがいかに自然言語を処理できるかを扱うコンピュータ科学の領域

ニューラル処理装置

NPU

フォンノイマン又はハーバードアーキテクチャでなく，ニューロモーフィングの原理に基づいた電気回路

整流リニアユニット

ReLU

二つの線形部分から成るニューロンの活性化機能

リカレントニューラルネットワーク

RNN

ニューロン間の接続がシーケンスに沿って有向グラフを形成するような，ニューラルネットワークのクラス

リレーショナルデータベース管理システム

RDBMS

リレーショナルモデルに基づくデータベースシステム

テンソル処理装置

TPU

機械学習とディープラーニングのタスクを処理するために Google が開発した特定用途向け集積回路

第 1 節

序文

1.1 人工知能：奇跡か、曇りか？

人工知能(AI)は現在、最も広く誇大宣伝されている技術の一つある。最初のコンピュータが出現して以来、より広範な意思決定プロセスにおいて人間を支援するため、数学モデルがますます使用されるようになってきた。ある仕事に誰を雇うかを定める助けとして人事の分野で使う場合であれ、銀行部門で融資に関して承認される受取人を選ぶために使う場合であれ、これまで人間の判断及び裁定に委ねられていた領域に、機械は継続的に入り込んでいる。

大量のデータの利用を可能とする多くの産業のデジタル化と共に、AIは増大する問題を解決するその可能性について新たな関心を集め始めた。特に人工ニューラルネットワーク(ANN)として知られているものとの関連で、機械学習技術はより強力で高度なものへ成長してきた。生物学に触発され、数学的好奇心として 20 世紀中盤に開発されたニューラルネットワークは、AIの礎石の一つとなった。

しかし、一般にディープラーニングと呼ばれている、機械学習の劇的な改善が AI の爆発的発展の道を開いたのは、2010 年以降になってからであった。計算能力が着実に増すにつれ、極めて大規模な("ディープ")ニューラルネットワークは、従来のプログラミング手法を用いて実行するには複雑すぎる、又は実行不可能でさえあった新しい機能をマシンに与え始めた。それ以来、コンピュータビジョン及び自然言語処理(NLP)などの技術は完全に変貌し、それらは様々な製品及びサービスにおいて大規模に採用

されている。ディープラーニングは現在、製造業、ヘルスケア又は金融など多くの産業で、新しいパターンを発見し、予測を行い、様々な重要な決定を導くために適用されている。

しかし、こうした最近の進歩は印象的なものであるが、AIは現在もなお非常にタスク指向であり、十分に定義されたパターン認識アプリケーションの周辺に集中している。状況認識又は感情移入のような人間的なスキルを機械に備えさせるために、現在の研究はダイナミックな作業を行っているが、多くの AI 科学者によれば、この目標の達成はまだ遠い未来のことである。

現在の限界にも関わらず、AIはすでに社会、ビジネス及び個人に対して深い影響を与えており、人々がいかに生活し、仕事し、互いに相互作用するかに対して影響を増していくことが予想される。あらゆる大きな技術シフトと同様、AIは偶像視されると同時に悪魔のように見られている。ロボットが次々に仕事を奪っていくというものから、特に人間に対して戦いを挑む AI 駆動の機械まで、AIがもたらし得るあらゆる種類の現実的脅威が考えられている。にもかかわらず、そのような暗いシナリオとは別に、新しい倫理的及び社会的な課題が、革新的な AI の発展と同時に出現していることは否定できない。ビジネス、政府、規制当局及び社会全体が、AIが全人類に真に恩恵をもたらすように、こうした課題に対処する必要がある。このような状況において、技術規格と適合性評価システムは、AIの将来を形作る上で重要な役割を果たし得る。

1.2 人工知能の冬の時代から再生まで

AI は、その起源を 1950 年代まで遡ることができるから、新しい科学分野ではない。文献は一般に、AI 開発の三つの歴史的段階を指摘している。

第一段階(1950 年代から 1980 年代)において、AI はプログラマブルデジタルコンピュータの抽象的な数学的推論から出現した。コンピュータの有名なバイオニア、Alan Turing は、あるプログラムをインテリジェントとみなし得るかどうかを決定する初めての試験、いわゆるチューリング試験[1]の概念を示した。実際には“人工知能”という用語は、後に AI の父の一人として知られるようになった John McCarthy によって 1955 年に初めて作られ[2]、1956 年にダートマスカレッジで開催された AI の最初の会議のタイトルとして提案された。

AI 開発の次の重要なステップは、Frank Rosenblatt による 1958 年の、ニューラルネットワーク(「パーセプトロン」)の概念を用いたアルゴリズムの発明であった[3]。しかし、機械学習が実際のアプリケーションで使用され始めたのは、Cover 及び Hart による 1967 年の近傍法アルゴリズムの開発[4]以降であった。これらの初期の達成及びコンピュータベースのシンボリズムの急速な発展にも関わらず、多くの概念を正式に表現又は描写することができなかったために、AI の到達範囲は限られていた。

第 2 段階(1980 年代～1990 年代後半)では、エキスパートシステムが急速に発展し、数学的モデリングで大きなブレイクスルーが達成された。ANN も、増え続けるアプリケーションに渡って、より広範に採用され始めた。この期間に AI の幾つかの中核的手法とアルゴリズム、すなわち、1981 年の説明に基づく学習(EBL)[5]、1986 年の逆伝播アルゴリズム[6]、及び 1995 年のサポートベクターマシン(SVM)の原理[7]が開発され、より洗練されたものになった。

この第 2 段階の最もよく知られたマイルストーンの一つは、1996 年に IBM が開発したディープブルーのチェスプログラムであり、これはその翌年、世界

チャンピオンを打ち負かした[8]。コンピュータプログラムが世界選手権レベルのゲームで人間のプレーヤーを負かすことができたのは、これが初めてであった。この成功にも関わらず、知識の獲得及び推論の能力に関わる制約は、AI システム導入のコストの高さと相まってある幻滅をもたらし、それがある観察者たちに“AI の冬”を語らせることになった。

AI が当初の見込みを達成し始めたのは、21 世紀初頭に始まった開発の第 3 段階においてであった。2006 年に、最初の強力な高速学習ディープビリーフネットワークが Hinton, Osindero 及び Teh の論文で紹介された[9]。このアルゴリズムは、一連の画像内の数字を認識及び分類するために使用された。この貢献は AI の開発に役立ち、現在の AI 研究に最も影響を及ぼした研究の一つとなった。2010 年の IBM Watson、及び 2016 年の AlphaGo のような最近の開発が、その後、一般の人々の大きな注目を集めた。

収集されるデータの爆発的増加、理論的アルゴリズムの持続的な革新、及び計算能力の継続的な向上と共に、AI はその後多くのアプリケーション分野でブレイクスルー的進歩を果たし、現在は新しい課題に取り組む十分な準備ができているように見える。こうした発展の全ては、あるアナリストたちに“AI の再生”を語らせることとなった。

図 1-1 に、初期から現在までの AI の主なマイルストーンの一部を示している。

音声及び画像認識のための機械学習アルゴリズムの成功は、研究コミュニティ、ビジネス及び政府から大きな関心を得ることに役立った。加えて、クラウドコンピューティングとビッグデータが並行的に開発されたことが、コンピュータベースの AI シミュレーションから、機械と人々を結び付けるより複雑でインテリジェントなシステムへの移行を支援した。AI は、第 4 次産業革命の中核的技術の一つになると共に、輸送、ヘルスケア、小売、教育、政府サービス、及びその他の産業の革新の駆動力になると現在予想されている。

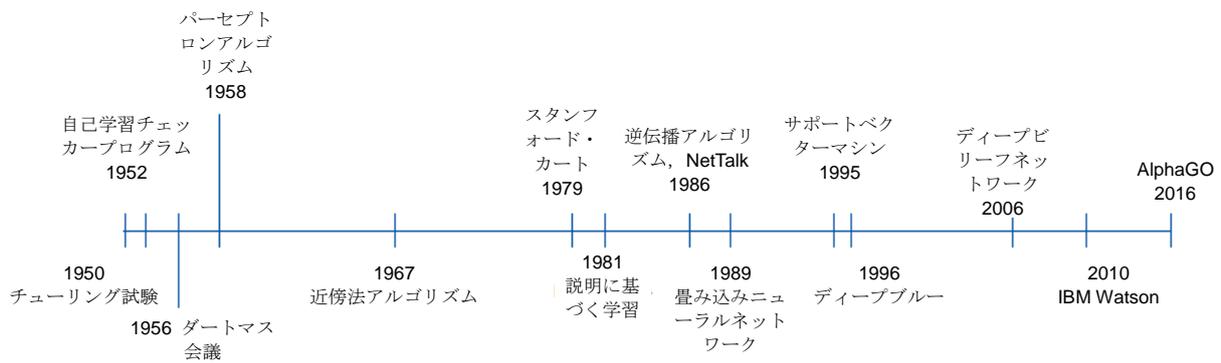


図 1-1 | AI 開発の主なマイルストーン

1.3 大きな機会はリスクと課題を伴う

AI は潜在的に巨大な市場をもっている。国際データコーポレーション(IDC)の最近の調査によれば、認知及び AI システムに関する全世界の支出は、2021 年に 570 億米ドルを超えると予想されている[10]。小売と銀行部門が、今後数年間において最も AI に支出をし、個々の製造、ヘルスケア及びプロセス自動化がこれに続くと予想される。これも IDC によれば、これら五つの産業は AI 技術の最大の利用者であり続け、それらの投資額を合計すると、こうした技術に関して、2021 年まで全世界の支出の約 55% を占める。

プログラム管理、教育、訓練、ハードウェア設置、システム統合及びコンサルティングを含む、マシンインテリジェンスの関連サービス産業を考慮に入れるなら、市場の規模は実際にははるかに大きく、AI は近い将来、最も急成長する産業の一つになると予想される。

自動化された顧客サービス及び診断システムは、今後数年間、AI に対する支出の最大の推進要因である可能性が高いが、一方、スマート製造が AI 市場で強力な地位を確立すると予想される。IDC は実際、インテリジェントなプロセスの自動化が、2021 年までに AI システムの 3 番目に大きな使用事例になると見ている[10]。急速な支出の増加が生じるその他の使用事例は、公共の安全、緊急時対応、及び買利物のアドバイザーや推奨などを含む。

これらの刺激的な市場見通しは、必然的に幾つかのリスクと課題ももたらすであろう。労働力に対する AI の影響が、社会への潜在的な脅威としてしばしば引用されており、雇用市場が徐々に多様化することの帰結として社会的関係に緊張が生じる。自動化と接続性が増すことで、先進国と発展途上国間の富の格差がさらに増す又は強まる可能性もある。そうしたシナリオは現在不確かではあるが、世界中のあらゆる主要経済は自らの戦略的な技術計画活動の一環として、AI の革新を支援する多額の投資を始めている。例えば中国は 2017 年に、“新世代の AI 開発計画”、“新世代の AI 産業促進のための 3 か年行動計画(2018~2020)”，並びに、AI 技術の研究、開発及び産業化を加速させるためのその他の様々な政策を公布した。

1.4 人工知能の定義

AI を定義する方法は幾つかある。ISO/IEC 合同専門委員会(JTC) 1 は定義を、“推論及び学習のような人間の知能に一般に関係する機能を果たすモデル及びシステムを扱う、通常はコンピュータ科学の一分野とみなされる学際的分野”であるとしている[11]。エッジインテリジェンスに関する IEC の白書[12]では、用語 AI は“パターンマッチング、学習、問題解決のような、人間が他の人間の心と関わる際の認知機能を、機械が模倣する”ときに適用される。

言い換えれば、インテリジェンスは四つの基本的能力、すなわち感知、理解、行動及び学習によって示される。理解は、現時点で人間の場合と機械の場合で同じ意味を持っていない。一般に、モデルは従来の方法に比べてより良く実行する方法を“学習”するように訓練されるであろうが、AI システムは未だ、自らの周囲世界を“理解している”と主張することはできない。

AI の実務者はしばしば、強い AI と弱い AI とを区別する。強い AI (一般的な AI と呼ばれる)は、人間のインテリジェンスを正確に模倣できる機械の、より哲学的な概念を指している。そうした機械は、先進的な認知能力が必要な全ての分野の全ての問題を解決できるであろう。この種の AI はまだ開発されておらず、様々な SF 小説又は映画の中でしか見ることができない。

これと対照的に、弱い AI (狭い AI と呼ばれる)は、特定の使用事例の特定の問題を解決する際に人間を支援する。例えば、AlphaGo はボードゲームである碁をほぼ完璧な程度までマスターするが、他の問題は何も解決することができない。Siri のような音声認識ツールは、異なった弱い AI を組み合わせた

一種のハイブリッドインテリジェンスを提供する。これらのツールは、話し言葉を翻訳して、様々なタスクを実行するために単語を自らのデータベースに結び付ける能力をもつ。にもかかわらず、こうしたシステムは、一般的な形態のインテリジェンスを構成してはいない[13]。

機械学習及びディープラーニングのようなその他の用語は、AI と密接に結び付いている。インテリジェントな機械を作るには、特定の種類の知識が必要とされる。かつては、そうした知識は直接機械中にハードコーディングされていたが、そのことは特定の制約と限界をもたらした。機械学習で採用されているアプローチは、所与のデータセットに基づき、機械が自分自身で知識を構築していくことである[14]。

最近では、知識はほとんどが現実世界のデータに由来するため、機械学習アルゴリズムの性能は、表現と呼ばれることもある利用可能な情報と密接に関係している。表現は、所与の機械に利用できる全ての特徴(例えば、予測保守アプリケーションでの温度又は振動センサの出力)から構成されている。正しい表現を選択することは複雑で時間がかかる作業であり、その分野に関する高度に専門的な知識が必要とされる。

表現学習と呼ばれる機械学習の分野は、生データからの特徴を検出又は分類するために必要な表現を発見することによって、この作業を自動化している。この一連の方法は、より伝統的な、特定のタスクを行うアルゴリズムとは対照的に、学習データの表現に基づいている[14]。

しかし、適切な特徴を選択することは、それらは様々な環境要因に依存しているため通常は非常に困難な作業である。例えば、暗い環境では色彩は異なって知覚され、そのことが対象のシルエットに影響する可能性がある。

表現学習のサブカテゴリであるディープラーニングは、特徴を変換して、受け取った入力に基づいて依存性を詳しく検討する。画像の例では入力特徴はピクセルである。ディープラーニングアプローチは、対象を識別するために、最初はピクセルを画像のエッジに、次に角に、最後に輪郭にマッピングする[14]。

図 1-2 は、これらの概念が互いに論理的にどのように関係しているかを示している。ディープラーニングは一種の表現学習であって、これは言い換えれば一種の機械学習であり、AI への全ての可能なアプローチの一つのサブカテゴリである。

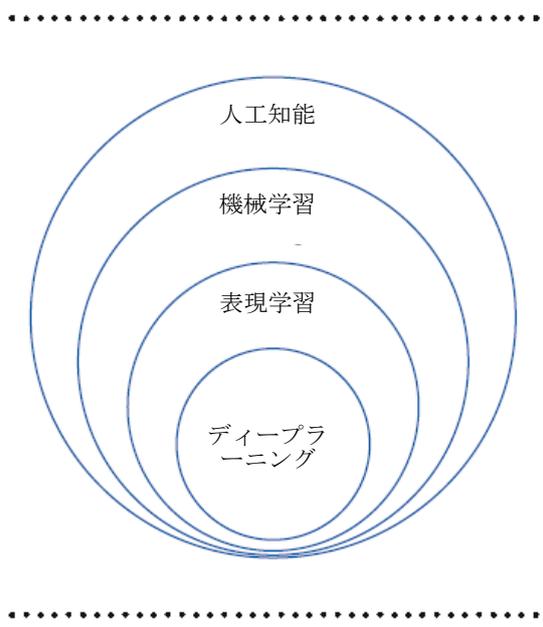


図 1-2 | AI のベン図

1.5 この白書の適用範囲

この白書の目的は、幾つかのアプリケーション領域について AI の機会と課題を探ることにより、AI が現在どの地点に立っているか、及び今後 10 年間の展望について概要を示すことである。幾つかの使用事例を基に白書は幾つかの推奨を行い、それらの一部は、IEC とその委員会において将来の標準化と適合性評価の作業のために対象とされている。

この白書は主に弱い AI に焦点を当てている。世界が現在、様々な形態の弱い AI を使用していることは明確である。Eメールのスパムフィルタは、古典的な種類の弱い AI である。これは、何がスパムを構成するか、何がしないかの判別を可能にする特定レベルのインテリジェンスを搭載して動作を開始し、個人ユーザー又はユーザーの集合の特定の好みを踏まえ、経験を得るに従ってそのインテリジェンスを洗練させていく。現在の自動車は、アンチロックブレーキをいつ作動させることが望ましいかを決定するコンピュータから、燃料噴射パラメータを調整する処理装置まで、弱い AI システムに満ちている。現在試験されている自動運転車は、車両が周囲環境を感知して応答することを可能にする多くのロバストな弱い AI システムを含んでいる。

人間の脳と同じくらいに賢い AI を作ることは、かなり長い期間に渡って大きな挑戦に留まるであろう。10桁の数字二つをごく短い時間で乗算できるコンピュータを構築することは、思いのほか容易である。犬を見て、それが実際に犬なのか猫なのかを決定できるコンピュータを構築することは、はるかに困難である。チェスの世界チャンピオンを打ち負かせるシステムの構築は既に達成されているが、6歳児の絵本から採った文章段落の意味を理解でき、それが単に単語を認識するだけではないような AI を作ることは、まだ現在の能力をはるかに超えている。

人間にとっては容易に見えるそうした能力の多くは、実際には非常に複雑であり、それに関係するスキルが、人間(及びほとんどの動物)にあって何億年もの進化の年月に渡り最適化されてきたため、容易に見えるにすぎない。インテリジェンスが何から構成されているかを適切に定義することは実質的に不可能であるので、強力な AI の開発で何を成功と考えるかについて、明確な基準を与えることは非常に難しい。

今後 10 年以内に予想される、幾つかの AI 開発に注目するため、この白書は次のアプリケーション領域について調べ、それぞれについて幾つかの使用事例を記述する：

- スマートホーム—温度, 照明, セキュリティ, 安全又は娯楽などの機能を制御, 自動化及び最適化するための接続製品を備えた住宅環境。
- スマート製造—製造プロセスとデータ分析を監視して製造業務を改善するために接続機械を使用する, 技術主導型のアプローチ。
- スマート輸送—交通の実状を改善し, 渋滞, 事故又は公害などの潜在的な悪影響を減少させるために, 移動に係る存在(例えば, 自家用車, 公共交通機関, 配達車両, 緊急サービス, 駐車場)を統合し, 自動化するシナリオ。
- スマートエネルギー—再生可能エネルギーの生産, 消費, 及びインフラストラクチャを, エネルギーサービス, それを利用しているユーザー, 及び情報通信技術(ICT)を通じて統合し, 調整する持続可能で費用対効果の高いエネルギーシステム。

1.6 この白書の概要

第2節では, 社会, ビジネス及び個人に大きな課題を課す幾つかの重要なメガトレンドを明確化することにより, AIの必要性を記述している。AIは, ほんの数例を挙げれば, 環境に関する懸念, 人口動態の傾向の変化, 又は経済格差のような課題の幾つかに対処する幅広いアプリケーションを可能にし, 向上させる。

AIは新しい分野とは言えないが, 劇的な技術の向上, 特に機械学習の分野でのそれがAIの現在の爆発的発展への道を開いたのは, 2010年以降であった。このブレークスルーは, 第3節で説明する幾つかの要因によって可能となった。主に, 計算能力の大きな向上, より高度な機械学習アルゴリズム, 及びAIシステムの訓練のために大量のデータが利用できたことが, 現在の目覚ましいAI発展の実現要因であった。

AI研究の分野の繁栄に寄与した, 幾つかのその他の推進要因について第3節でさらに概要を記述してい

る。それらは, クラウド及びエッジコンピューティング, モノのインターネット(IoT), 及びビッグデータのような情報技術(IT)の開発, 並びに, 新しい技術を利用しデータを共有することに対して消費者と社会の準備が以前よりできたことを含んでいる。

第4節は, 最も一般的なAIシステムと機械学習技術についての高度な理解を提供する。技術的な詳細に深く入り込むことなく, 将来のAI開発の基礎となる現在使用されている最も人気のあるAIアルゴリズムについても概要を記述している。AIの現在の先端状況に基づき, ここに示す技術の概要は, AIが実際に内部からどのように機能しているかに関して確かな科学的な理解を得たい読者のために, 多くの参考文献により補足されている。

この白書は, 考え得る全てのAIアプリケーションのシナリオを扱うことはできないが, 最も一般的なアプリケーション分野の幾つかに対して, 現在の主なAIシステムがどのように対応しているかを第5節で説明している。この記述は, 幾つかの産業部門におけるAIのニーズと要求事項の特徴をより良く示すものになっている。残りの部分では, 四つのアプリケーション分野(スマートホーム, スマート製造, スマート輸送, 及びスマートエネルギー)をより詳しく説明しており, これらの分野に関しては, 幾つかのAI関連の使用事例とシナリオを, 現在の及び新たなAI利用の最も急を要する課題の幾つかと共に, レビューしている。

このレビューに続き, 第6節では, 現在の利用において特定し得る, 又は新たなAI開発において予測され得る主なAIの課題を整理している。課題は次のような幾つかのカテゴリに分けられる。すなわち, 社会的及び経済的課題; 訓練データの選択及びデータの標準化を含むデータ関連の課題, アルゴリズムのロバストネス及び解釈可能性のようなアルゴリズム関連の課題, ハードウェア又はプラットフォームに関連するインフラストラクチャの課題, 信頼, プライバシー及びセキュリティを含む信頼性の課題, 並びに責任及び倫理的事項のような規制関連の課題である。

前節に基づいて、第 7 節では AI の標準化の展望を示し、前に記述した AI の課題の一部を解決する上で対処する必要がある幾つかのギャップを明確化している。現在の標準化の取り組みはまだほんの初期段階であるが、この白書は AI の将来を形作り、諸産業での AI の採用と共に生じる多くの技術的及び非技術的問題を軽減する上で、規格が本質的な役割を果たす必要があることを明確に示している。

最後に、第 8 節では産業、規制機関及び IEC に対する一連の推奨事項を提示して、この白書を締めくくっている。多様な社会における AI の成功とその受容は、複数のステークホルダーとコミュニティの関与に依存しているが、AI がその見込みを確実に実現する上で、産業、政策立案者、及び IEC のような規格策定組織がこれを推進する役割を果たす必要があることはこの白書から明確である。

今後 10 年間で重要性が増す可能性のある、将来を展望したその他の AI 開発について附属書 A で考察している。

P20

第2節

人工知能のニーズ

現在の社会及びビジネス環境は、複雑で前例のない一連の課題と機会によって特徴付けられている。既存の市場は混乱にさらされており、短期間で突然消え去る可能性さえある。しばしばメガトレンドと呼ばれる、社会、経済、ビジネス、文化及び個人の生活に影響を与える大きな世界的トレンドが、人間世界の将来、及び増しつあるその変化の速度を決定付けている。

メガトレンドは、AIのような大きな技術発展の影響の形成に寄与する地球規模の相互作用を表している。デジタル化、自動化及びAIの組み合わせの効果は、仕事の将来に大きな影響を与えると予想される。コンピュータ化は熟練性が低い多くの仕事に影響を与え、コンピュータを利用した自動化は、製造、計画及び意思決定を含む多くの産業と環境において、ますます普及していくと予想される[15]。技術力の向上は既にサプライチェーンを変化させ、労働力の形を変え、仕事を再定義している。こうした変化の見通しが課題を発生させるのは、成長が直線的でなく複雑で加速的であるという事実ゆえである。

同時にAIは、これらのメガトレンド、そのほんの数例を挙げれば、環境に関する懸念、人口動態の変化、又は経済格差などから生じる、幾つかの課題に対処できる幅広いアプリケーションを可能にし、改善するであろう。

2.1 天然資源の欠乏

地球の天然資源は驚くべき速度で消費されており、ほとんどの国が2050年までにそうした資源の年間世界消費量を倍増させることが予想されている[16]。限りある天然資源が枯渇しつつあるだけでなく、自

然が再生できる量よりはるかに多くの環境資源を人間は使用している。過去には資源の保護はしばしば、ビジネスに不利益であるとみなされていたが、現在、この二つは決して相互に排他的ではない。

AIは既に、無数の製造業者たちが生産プロセスを最適化する助けになっており、それによって廃棄物を減らし、生産を増大させている。さらにAIは近い将来、プロセス自体だけでなく、それらのインプットの最適化にも使用されるであろう。製造の投入材料の目的、特性及び環境影響を分析することによって、AIは、より持続可能な製造で求められる仕様に合った材料を科学者が設計することに役立ち得る。AIを、機械が作り出す副産物的な材料成分の二次的使用法を明確にするために使用し、それによって原材料を循環使用に近い形で使うアイデアさえ提案されている。製造プロセスにおけるこうした効率化の恩恵は、ビジネスに対して魅力的なインセンティブになるだけでなく、世界の資源消費にも大きな影響を与える。

AIは、都市化が進み、電力消費が増し、水資源が不足し、再生可能エネルギーが大規模で配備される時代において、公益事業でも役立つ。これは需要と供給の両方をよりインテリジェントに管理することで達成される。需要サイドではAIは既に、例えばデータセンターのエネルギー消費を15%減らすことによって、大きな省エネルギーをもたらしている。供給サイドでは分散型スマートエネルギーグリッドにより、化石燃料の使用を最小限に抑えつつ最適な供給レベルを確保するため、停電を予測してそれを回避し、供給と需要の変動を管理することができる[17]。

その他の例は、太陽光発電所又は風力発電所による再生可能エネルギー発電の最適化、あるいは排気ガスを減らすための自動車交通量の最適化などを含んでいる。

2.2 気候変動

気候変動問題に対する一貫した対応と、より責任ある環境資源の使用がなければ、予測のできない変化が地球を脅威にさらすという点で、先導的科学家たちの予測ははっきりしている。このことは、経済的な目標と環境の持続可能性をいかに調和させるかという問いを提起する。同様に重要なのは、人類がいかに将来の予想外の劇的な自然事象に対し、準備できるかということである。

多様なアプリケーションでの AI の使用は、気候変動との闘いにおいて主導的な役割を果たすことが期待されている。AI は、天然資源を扱うとき又は想定外の出来事を予測するときに、複雑な意思決定プロセスを支援できる。例えば、資源の消費と使用は、発電の際に既に最適に調整されている可能性がある。AI は、天気予報及び電気料金との関係で、状況に応じた電力消費及びグリッド負荷のような多くのパラメータを設定することを可能にする。その結果、電力の消費者の挙動をより効率的に判断し、それに対処できる。

これらの達成に基づいて、自律的エレクトロモビリティを含む、より責任ある資源使用に関わるインテリジェントモビリティ・ソリューションを採用することができる。乗用車とトラックを互いに最適かつ効率的にマッチングできるだけでなく、AI のおかげでそれらをより効率的に運転できる。

同様の開発を、効率的な水消費に関して考えることができる。例えば農業では、個々の植物が必要とする特定のもの、土壌の状況、及び現在の気象条件に応じて最適な水需要を決定することを AI は可能にする。さらに、干ばつ及び水不足に関し、それに影響される地域と国々における供給の戦略を、AI を使

用して策定できる。

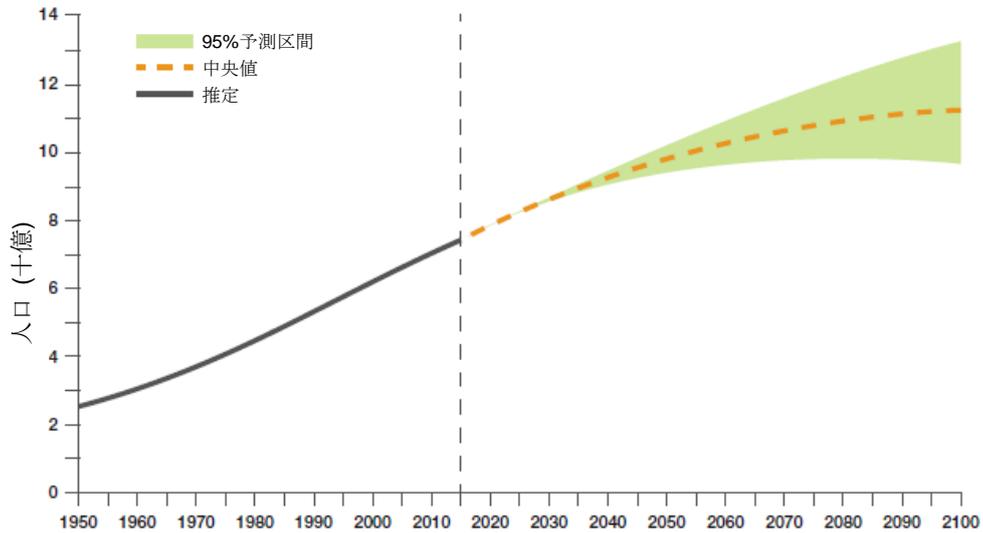
AI は、気象のシナリオ及び自然災害の予測向上にも貢献できる。科学者は、天気予報の精度を高めるために、多くの影響因子をとらえて処理するという課題にいつそう直面するようになっている。AI は、天気に関連する事象を早期に予測し、洪水、大気汚染の兆し、又は暴風雨のような潜在的要素について警告するために、人々が広範な測定データを処理することを効果的に支援する。その後、早期警報システムを、多様な地域に対してよりインテリジェントに設定することができる。

2.3 人口動態の傾向

国連(UN)は、新興国及び開発途上国の人口増加により、2030年までに10億人を超える人口増加を予測している(図2-1)。さらに、高齢者人口(65歳以上の人の数)は、人がより長生きし、子供がより少なくなっているため、3億9000万人以上増加する。

図2-2に示すように、高齢化の影響は、欧州、アジア及びラテンアメリカでより早期に感じとられており、地域によって異なる問題になっている。1人の高齢者を支える、働く人々の数の傾向は、アジアでは2015年の9人から、1/2人~4人の範囲に移行し、このことは若い世代への高齢世代の依存性をより大きく高める。欧州では、適切な労働年齢人口が得られなくなって行き、新世代の労働者、すなわち女性と高齢者の需要が急激に生じる。高齢者1人当たりの労働年齢人口の割合は、2015年の4人から、2050年には50%減少する。これらの高齢化の傾向は、政府と業界にとって大きな課題となる[18]。

こうした傾向の別の重要な側面は、労働年齢人口が得られる点に関する大きな地域差に関連しており、例えば、2050年の日本では高齢者1人当たりの労働年齢人口が1.55であるのに対し、ナイジェリアでは高齢者1人当たりの労働年齢人口は15である。



出典： 国連, 経済社会局, 人口部(2017)
世界人口見通し：2017年改訂。ニューヨーク：国連

図 2-1 | 増加する世界人口

世界は高齢化している

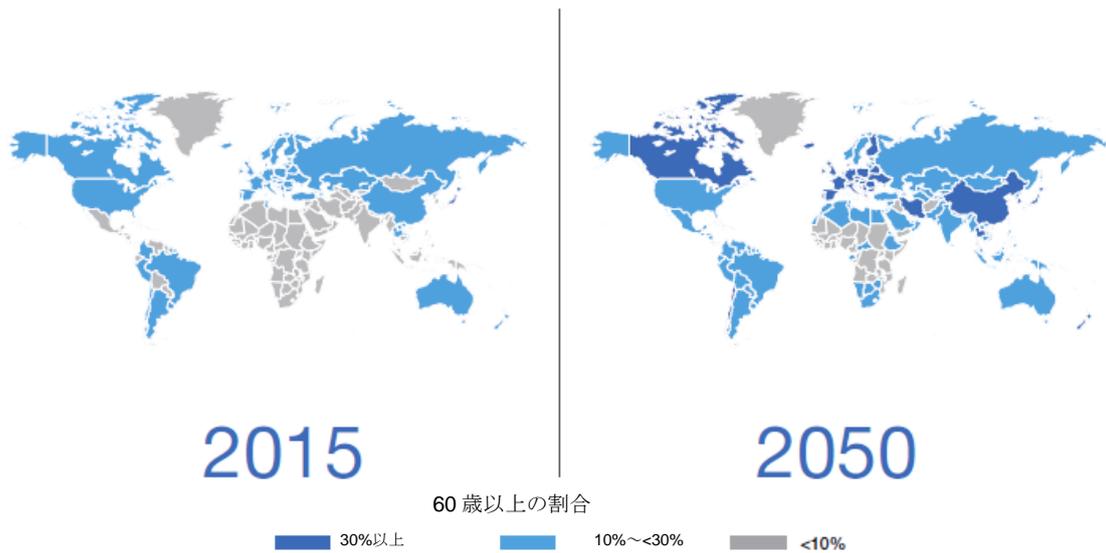


図 2-2 | 世界における高齢者人口

人口の高齢化傾向は、ヘルスケアへの支出を増加させる。G7 諸国では医療予算が毎年、約 2,000 億米ドル増加すると予想されている。しかし、技術的及び科学的な革新が、医療費をより対応可能なレベルへ減少させる助けとなり得る。

欧州のほとんどの政府は、現在、公式の退職年齢を引き上げ、年齢による差別を違法とすることによって、高齢の労働者がそのまま労働力であり続けることを奨励している。さらに、産業界は高齢の労働者のために、金銭的インセンティブと再訓練プログラムを提供する必要がある。各人の仕事生活の中で、新たなスキルを獲得するための生涯学習、及び若い同僚に対する指導が、高齢者が労働力であり続けるようにする取り組みのために常に重要となる。

2.4 経済政策

自動化と生産性の著しい向上が、人口動態の変化のプレッシャーを諸国が緩和することにより役に立つかは、既に注目されている事柄である。これは、高齢化に苦しむ先進的経済にとって明らかに歓迎する発展である一方、世界の最も貧しい国々に対して AI が変化をもたらす無数の機会が存在している。

例えば、国連の持続可能な開発目標は、貧困と飢餓を減らし、教育を改善することによってこれらの課題に取り組もうとしている。最近の“AI for Good Global Summit 2017”は、これらの取り組みをどのように AI が支援できるかに注目していた。これらの目標を達成し、最も資源が必要とされている場所を決定することに向けて、国際社会の進歩を監視するためのイニシアチブから、病気発生の予測モデリングまでに、諸提案は渡っていた[19]。AI は、スマート農業推進のために IoT、ドローン、及び合成生物学と連携して作動したり、作物をいつでもどこで植えて収穫すべきかについて洞察を提供しつつ、収穫を増やして世界の飢餓と闘うことを助けるために栄養分、農薬及び水を最適化したりすることもできる[20]。

AI の可能性に対して政府の関心が急速に高まっており、そうなるにつれ、公共政策、特に経済政策に対して AI が貢献することが期待されている。エコノミストは通常、モデルを構築するときは、その後検証しようとする一連の仮定から出発する。しかし、AI はデータを分析し、変数の間の未知の相互作用を発見する力を与える。この相互作用は、それに基づいてモデルを第一原則から構築し、公共政策と金融政策に対して情報を提供できるようなものである。

AI は、健全性又は行為の観点から懸念される異常事項について、銀行の貸借対照表を検査することにより、金融規制機関の監視作業も支援し得る[21]。政府が最良実施法を探求し、より多くの情報に基づいて AI 政策を決定し、公衆の信頼を築くことに役立ち得るので、公的領域でどのように AI を使えるかに関する探求は歓迎することが望ましい。

2.5 サービスと製品のカスタマイズ

AI と先進的な製造を統合することによって、供給業者、パートナー及び顧客を容易に結び付ける大量のカスタマイズが可能になり、大量生産に近い効率とコストで個人個人の需要を満たせる。このように、AI を適用すると製造のバリューチェーンが最適化され、製造業者は材料の流れをリアルタイムで追跡できる。エンジニアリングと品質の問題をより正確にアセスメントし、過剰な在庫と物流の遅延を減らし、顧客のニーズへの応答性を高め、廃棄物とコストを削減するより良いビジネス決定を行うこともできる。インターネットに接続している消費者が、自らが望む仕様によって製品を開発するためインテリジェントな製造プロセスをコントロールできるようにすることで、ビジネスは生産の大量カスタマイズから恩恵を得る。

AIは、全く新しいレベルでの製品の個別化を可能にする。顧客がオンラインで使用できる製品構成だけがこれに影響されるわけではない。AIの使用は個別化の全く新しい可能性を開く。個別化に高い開発コストを必要とする製品は、AIを使用して要求事項に適合させることができる。

サービスも、顧客のニーズに合わせて自動的に調整できる。一例は、文章の自動翻訳である。従来、規則に基づいた手法の結果は、しばしば、文脈を作っている文の全体ではなく個々の単語の意味を反映できるだけだったが、AIが支援するサービスは文章の意味に基づいて翻訳を実行できる。

第3節

人工知能の可能の実現要因と推進要因

ハードウェア、アルゴリズム及びデータの利用可能性の向上が AI の主な実現要因であり、資本はその燃料であった。新しい革新に対し、資金と市場の両方を与えることに寄与した AI に関する意識の高まり、ベンチャーキャピタル、及び政府の支援がなければ、現在見られる急速な発展は恐らく可能でなかった。AI を効果的に使用できる人々に対して AI が与え得る利点について、認識が高まってきている。その結果、AI の将来を形作り、それによって AI 開発のより良い市場条件を作る上で、ビジネス及び技術のリーダーたちがはるかに積極的な役割を果たすようになっている。

• 意識の向上

2017年に調査を受けた組織の78%が、将来、AIを採用する計画があると報告し、50%近くが採用について積極的に検討していると回答している[22]。実際、2016年には、諸企業が合計でAIに390億米ドルを投資し[23]、特に機械学習に対するものが多く、それが投資の60%近くを集めていた。中小企業(SME)によるAI技術への関心、及びその採用の増加が、今後長年に渡ってこの成長の燃料となり続けることは疑いがない[24]。したがって、AIアプリケーションの潜在的な市場は巨大であり、2025年までに総市場規模は1,270億米ドルに達すると予測されている[23]。

• 民間資本の利用可能性

諸企業がAIを通じて自らのビジネスを変える意欲をもち、その対価を支払う意志がある中で、AI起業家にとって資本の利用可能性はかつてなく高まっている。世界のベンチャーキャピタルのAIへの投資は、2017年に倍増して120億米ドルになり[25]、AIの新興企業の数、米国だけでも2000年以来14

倍に増えている。技術関連企業も、自らのAI部門への10億ドル規模の投資を発表して互いにしのぎを削っている。

• 政府の支援

こうした民間資本の利用可能性は、政府が自国のAI産業を成長させることを競っているのが増加だけが予想される状況である。中国が、2030年までにAIの世界的リーダーとして米国に追いつくという意欲的な計画を発表した後、欧州連合、英国、ドイツ、フランス及びカナダが、いずれも自国のAI産業を強化することを表明した[26]。表明された投資額は国によって大きく異なっているが、政府の関心の高まりは、官/民パートナーシップ、及び公共部門のためのAIアプリケーションの開発の基盤を築きつつある。

現在、AIの世界へのアクセスすることはかつてなく容易になっている。フレームワーク、ツールキット及びライブラリは、アルゴリズムと様々なプログラミング言語をユーザーに提供している。それらは、こうしたアルゴリズムを維持し、その実施も容易にし、開発者とユーザーのコミュニティを、共同でオープンソースソフトウェアを改良することへ導いている[27]。

3.1 人工知能の実現要因

AIへの関心は新しいピークに達している。ほぼ毎週、新しい発見が起き、新しいアプリケーションの広告がなされている。過去には考えられなかった達成が、現在、ペースを上げてなされている。例えば、IBM Watsonは、高い言語スキルと、なぞなぞ及び言葉遊びに関する知識を必要とするゲーム Jeopardy で、最高のプレーヤーを打ち負かした[28]。

残る問いは、近年、AIの進歩でこうした途方もない加速を可能にしたものは何かということである。AIの中核的なコンセプトとアプローチは数十年前から存在していたが、下記の三つの重要な実現要因が、“AIの冬”を回避することを助け、現在の壮観な発展をもたらした：

- 計算能力の向上
- データの利用可能性
- 改良されたアルゴリズム

3.1.1 計算能力の向上

ほとんどのAIアルゴリズムは、特に訓練の局面において極めて大きな計算能力を必要とする。計算能力が高いことは、アルゴリズムをより速く試験し、訓練できること、及びより複雑なアルゴリズムを実行できることを意味する。したがって、AIの採用の増加は多分にハードウェア技術(例えば、集積回路、半導体製造、サーバ技術)においてなされた進歩によ

って可能になった。

計算能力の向上は、こうした能力をチップ上のトランジスタの密度と関係付けるムーアの法則によって一般的に表される[29]。ムーアの法則に従って、半導体の形状サイズは1970年代の10 μ mから2017年の10nmへ縮小し、これは、同じダイサイズ上にはるかに多くのトランジスタを集積できることを意味する(図3-1)。

トランジスタ数が著しく増加して、はるかに高い計算能力が得られただけでなく、ハードウェアアーキテクチャも改良され、AIアプリケーションのパフォーマンスが向上した。例えば、並列性を高めるためにマルチコアプロセッサが設計された。中央処理装置(CPU)に加え、グラフィックス処理装置(GPU)、フィールドプログラマブルゲートアレイ(FPGA)、及び特定用途向け集積回路(ASIC)のようなその他の種類の処理装置が、様々な作業負荷パターンのためにさらに採用されつつある。

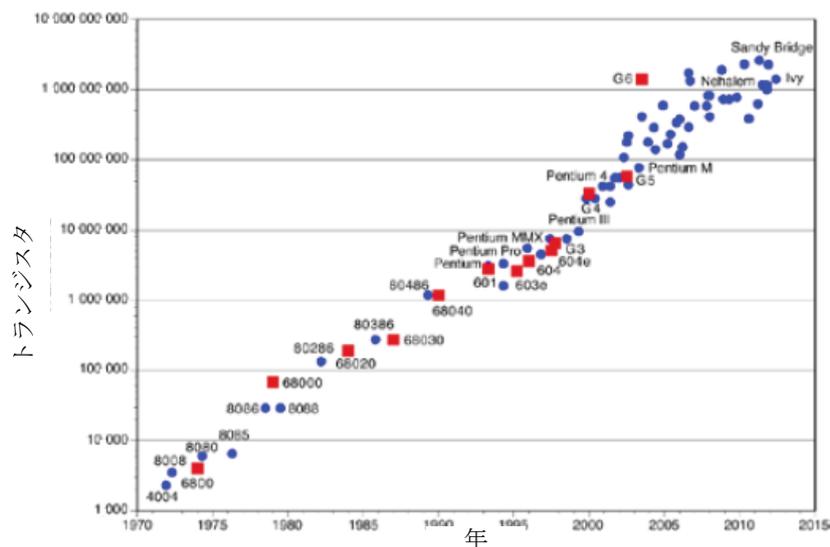


図 3-1 | 時間に対する計算能力の向上

GPU は画像処理タスクを扱うために使われており、ディープニューラルネットワーク(DNN)又は畳み込みニューラルネットワーク(CNN)のような AI アルゴリズムを加速する上で、非常に効果的であることがわかっている。FPGA のような集積回路は、製造後に広範なアプリケーションに適合するように構成することができ、従来の CPU に比べてはるかに高速である。ASIC は、テンソル処理装置(TPU)及びニューラル処理装置(NPU)のような様々な変種を含んでいる。データ処理パターンに基づいて回路設計を調整することによって、ASIC は GPU 又は FPGA よりも高い性能を達成できる。例えば、TPU は現在の CPU 及び GPU よりも、ワットあたり 15~30 倍、及びワットあたり 30~80 倍高い性能を提供すると言われている[30]。

3.1.2 データの利用可能性

AI アプリケーションの出力は、供給されたデータに基づいてアルゴリズムから抽出された情報である。したがって、不完全な又は誤ったデータを使用すると、アルゴリズムがどれほど優れていても、良い結果は決して得られない[29]。新しいアルゴリズムを作り、評価する際の重要な要因は、既に分類がなされているデータセット及び意味のあるデータにアクセスすることである。データの利用可能性を促進した最も重要な発展の一つは、インターネットである。これは、世界中の研究者がアクセス可能なデータセットを作るために、巨大なコミュニティが協力することを可能にした。

画像分類のために常にデータを作り、分類し、ラベル付けし、アップロードするインターネットコミュニティの一つの説明的事例は、ImageNet コミュニティである。訓練データを作り、ラベル付けすることは、以前は多大な時間がかかっただけでなく、ニューラルネットワークの訓練のために必要とされる大きなデータセットの場合はほぼ不可能であった。1997 年に、顔認識のための画像データセットは約 165 インスタンスから構成されていたが、顔の ImageNet データセットは 1,570 の完全に分類されたインスタンスから既に構成されている[32]。

ImageNet だけで合計約 1500 万の分類された画像を提供しており、これは容易にアクセスでき、AI アルゴリズムの訓練と評価に使用することができる[33]。

アルゴリズムの訓練と微調整に使用できる大量のデータに容易にアクセスできることは、AI の研究者及び実務者が新しいアルゴリズムの改善と開発に専ら時間を使うことができ、その後速やかにそれらを試験及び検証できることを意味する。これは、ほんの 20 年前には不可能なことであった。

3.1.3 改良されたアルゴリズム

AI の最近の進歩のほとんどは、ディープラーニングにおいて起きている。このコンセプトは数年前から存在していたが、実際のブレークスルーはごく最近起きた。これらはニューラルネットワーク研究において近年の大きなマイルストーンではなかったが、開発は停止しなかった。既存の技術の多数の改良と幾つかの新しい技術の開発が、多くのニューラルネットワーク実施の成功をもたらした。

ニューラルネットワークが情報をより速く、より効率的に処理することを可能にしたアルゴリズムの小さな変化の一例は、整流リニアユニット(ReLU)のアクティベーション機能である。アクティベーション機能は、ニューラルネットワークを構成している全てのニューロンにとって不可欠な部分である。ReLU のコンセプトは 2000 年に Hahnloser によって導入された[34]。これはランプ関数の形をとっており、二つの線形部分から構成されている。その実施の成功が最初に実証されたのは 2011 年で、それがニューラルネットワークをより効率的に訓練するために使用されたときであった[35]。これは別のアクティベーション関数、すなわちロジスティック関数又はシグモイド関数の代替となるものであり、線形関数よりも多くの処理能力が必要な対数曲線を記述する。ReLU のアクティベーション機能は、シグモイドのアクティベーション機能に比べ、効率的な伝播、スケール不変性、及びより高い計算効率のような多くの利点を提供する。これはディープネットワークにおいてより速やかで効率的な訓練を可能とし、複雑なデータセットを含むアプリケーションでは特に有益である。

ディープラーニングの発展に貢献した新しいコンセプトの別の例は CNN で、これは 1989 年に最初に導入された[36][37]。これは画像認識タスクに広く使用されており、そこでは人間を超えるパフォーマンスが達成できる[38]。

これら全ての開発によって、AI 研究はここ数年の間に極めて大きな進歩を遂げることができた。利用可能なデータ量の増加は、適切なアルゴリズムでそれを効率的に処理する能力がなければ有用ではなかった。AI の進歩が、単一の実現要因の結果ではなく、むしろ時間と共に徐々に改善されてきた様々なアイデアと技術の組み合わせの結果であることは明らかである。

3.2 人工知能の推進要因

上記の三つの重要な実現要因が、現在見られる AI 研究の繁栄分野を生み出した。知られている AI の経済的価値に加え、これらの開発も、幅広いアプリケーションに渡って AI の採用を加速させてきた多くの技術と社会的推進要因のおかげで可能となった。これらは全て、過去 10 年間に渡って産業、社会及び個人の生活にも浸透してきたデジタル変革のトレンドに関連している。

IT 開発は、AI アプリケーションの採用を支援する上で役立ってきた。クラウドコンピューティング、ビッグデータ及び IoT を含む、しかし、これらに限定されない現代の IT インフラストラクチャによって、豊富なリソースが提供されている。

クラウドコンピューティングは弾力的でスケラブルなインフラストラクチャであり、これは最小限の管理の労力でオンデマンドで提供できる、リソース及び高レベルのサービスの共有プールに対するアクセスを与える。AI の計算ニーズは、データセットとアルゴリズムにより大きく異なるため、クラウドインフラストラクチャが提供する向上したリソース使用率と効率を活かすことで、特定のアプリケーションの要求事項を満たすことができる。

ビッグデータは、分量、速度、多様性及び/又はばらつきにより特徴付けられる大きなデータセットを保存、管理及び分析する科学と工学である[39]。構造化データ、半構造化データ、及び非構造化データを扱う手法とアーキテクチャが提案されている。これらは例えば、リレーショナルデータベース管理システム(RDBMS)、分散ファイルシステム、グラフデータベース、及びこれらのデータを処理又は分析するための様々な計算フレームワークを含んでいる。

最後に重要なことであるが、IoT は物理的な装置、車両又はその他の電気機器のネットワークであって、これはそうした対象が互いに通信し、相互作用することを可能にする。IoT は、地理的に広く配置されたセンサを通じて、これらの装置の状況に関連したデータを収集するインフラストラクチャを提供する。これにより、アクチュエータを通じて装置を設定及び制御できる。IoT インフラストラクチャと AI 技術の組み合わせは、スマート製造、スマートホーム、及びインテリジェント交通のような多くのアプリケーションをもたらした。

これらの IT 開発は、ソーシャルメディアのようなデータ集約型ツールの受容及び広範な利用を促した社会変化と結び付いている。このトレンドは、AI の広範な採用を推進及び促進する別の要因であった。

3.2.1 クラウド及びエッジコンピューティング

エッジコンピューティングは、ネットワーク、ストレージ及びアプリケーションの機能を統合した、関連する物又はデータソースに近いネットワークエッジにある分散型オープンプラットフォームである。モバイル装置又はセンサの近傍で動作することにより、エッジコンピューティングは集中型クラウドノードを補完し、データの発生場所及び消費場所の近くでの分析と情報生成を可能にする。これにより、機敏な接続性、リアルタイムサービス、データの最適化、アプリケーションインテリジェンス、セキュリティ及びプライバシー保護に関して、産業デジタル化の重要な要求事項を満たすことができる。クラウド及びエッジコンピューティングは、費用対効果に優れたスケラブルなコンピューティング

リソース及び特殊サービスへのアクセスを可能にする。

AI は、次のような形でエッジコンピューティングの利用から恩恵を受ける：

- データの取得と保存をローカルに行うことで、データの前処理が可能になり、生データではなく、意思決定又は警報だけがクラウドサーバに送信される。
- エッジデバイスに機械学習アルゴリズムを入れることによって、より迅速で効率的な意思決定が行えるため、クラウドサーバにコンタクトする頻度が減り、意思決定に関する往復遅延の影響が恒常的に減少する。
- ローカルな識別管理とアプリケーション固有のアクセス方針を使用して、地域の規制に従ってデータをデータソースの近くでセキュリティ保護することができる。
- エッジコンピューティングのノード間の通信によって、分散した AI ノード間における AI 機能の分散とインテリジェンス共有が可能になる。

エッジコンピューティングとエッジインテリジェンスについての詳細情報は、エッジインテリジェンスに関する IEC の白書[12]に記載されている。

3.2.2 モノのインターネット

IoT は、デバイスからデータ、特に生産及び消費者情報に関するデータを収集することに焦点を当てている。過去数十年に渡っての ICT の発展は、より小さく、よりスマートなデバイスへのコンピューティング機能の広がりをもたらした[40]。こうした開発に基づいて、IoT は ISO/IEC では“相互接続されたオブジェクト、人々、システム及び情報リソースのインフラストラクチャと、それらが物理世界及び仮想世界の情報を処理し、応答できるようにするインテリジェントサービスを合わせたもの”と定義されている[41]。

全世界に設置されている接続機器の数は、2018 年の 230 億超え程度から、2025 年には約 750 億に増加すると予測されている[42][43]。これはデータ取得に対する IoT のインパクトを示している。将来は AI アプリケーションで使用できるデータの量がますます増えて、アルゴリズムのパフォーマンスを向上させるであろう。

現在、IoT のアプリケーションが、デバイスに接続されたセンサを使用し、パフォーマンスと環境関連のデータの取得を可能にしている。これはローカル又はクラウドプラットフォームを通じたデータ解析を可能にする。将来、こうしたシステムのリアルタイム動作は、自律運転のような時間が肝心のアプリケーションでますます重要になるであろう[44]。

接続されたこれらのセンサが生成する大量のデータを考えると、エッジにおける AI コンピューティング(すなわちエッジコンピューティング)の役割はさらに重要になる。前述のように、生成された全てのデータを中央へに送信し、それらのデータを分析し、その後デバイスに必要な情報を送り返すことは現実的でなく、ときには不可能でさえある。したがって、ローカルで簡単な分析又は決定を行えることが極めて重要になる。この傾向は、エッジコンピューティングに依存するデバイスの上でローカルに作動するような、より単純な AI アルゴリズムにも帰結するであろう。エッジの AI は、IoT を次の機能レベルに導くであろう。

IoT についての詳細情報は、スマートで安全な IoT プラットフォームに関する IEC の白書に記載されている(IoT 2020)[44]。

3.2.3 ビッグデータ

全世界に渡る接続性とネットワーキングの先例のない成長が大量のデータを生み出しており、この生成の速度は加速している。ビッグデータは、様々な手法とアーキテクチャを通じて、これらの非常に大きなデータセットの保存、管理及び分析を可能にする。ビッグデータは個人と組織にとって非常に有益であり、よりスマートな都市、より迅速な医学的ブレイクスルー、より効率的なリソースの使用、及び人的資源プロセスなど、多くの分野に対して洞察を提供し得る。

多くの組織がこのパラダイムを採用しており、意思決定、製品とサービスの開発、さらには顧客、従業員、供給業者、その他のステークホルダーとやり取りする際に、よりデータに基づいて行動するようになってきている。ソーシャルメディア・プラットフォームは、マーケティング担当者が現在どのように自らの顧客にアプローチし、マーケティング分野を変化させているかを示す良い例である[15]。

ビッグデータは一般に、「情報の取得、保存、配布、管理及び分析を可能にするために先進的な手法と技術を必要とする、大量の高速、複雑で変化し得るデータを表す用語」と定義されている[45]。文献[39]によれば、ビッグデータは次のように、量(volume)、多様性(variety)、速度(velocity)、及び正確さ(veracity) (4つのV)で表すことができる：

- 量—生成されるデータの量。この量は、IoTを通じて接続されるデータソース数の増加、それらの解像度の増加、及びデータの深さのために爆発的に増えている。AIアプリケーションの課題は、非常に大量なこのデータを処理、分析及び維持することである。
- 多様性—データソースの多様性に起因するデータの不均一性。複数のソースが一つのイベントを記述し、異なったデータフォーマットを構造化された形で、又は構造化されていない形でさえ提供している。これらのデータはセンサのデータに限定されず、例えば機械オペレータの専門知識のようなものでもあり得る。そのため、AIは異なったデータ種類の異なったソースからの情報を利用する必要がある。
- 速度—データが生成される速度であり、これは現在、多くの場合リアルタイムである。ある種のアプリケーションでは、これはデータの有効性の条件になるので、データの生成速度は重要である。これはしばしば、データの生成速度とその処理速度の間のトレードオフをもたらす。生成と処理の間の待ち時間は、AIアプリケーションに関して重要な要素である。

- 正確さ—データの質。上述のように、AIアルゴリズムの力は供給されるデータの力と同程度であるにすぎない。質の低いデータに基づくアプリケーションは誤った予測をもたらすので、AIは有用な結果を出し続けるためにデータの質の問題を減らす必要がある。

3.2.4 消費者の受容

AIのもう一つの推進要因は、新しいテクノロジーを採用し、データと情報を共有し、共同作業のコミュニティに参加してAIアプリケーションを向上させることに対して、消費者と社会全体の準備がますます進んでいくことである。

コンピュータ、スマートフォン、及びその他の電子機器と共に育った“デジタルネイティブ”と呼ばれる世代は、新しい技術を採用し、個人データを共有することに既に非常にオープンマインドになっている。現在、データのプライバシーに関する懸念が注目を集めるようになってきているが、若い世代は既に自らの生活の一部としてデータ集約型の行動をしている(ソーシャルメディアなど)。Deloitteによる最近の研究では、諸企業と情報を共有しようとする意志は、2014に比べて2倍に増えている。調査した人々のほぼ80%が、そのことで直接的な利益があるなら自らの個人データを共有する意思があると述べている[46]。これが、なぜソーシャルメディアがAIの主な適用分野であるかを説明する理由の一つである。

ソーシャルメディア・プラットフォームは急速に、個人、組織及びビジネスに関する共有、通信、ネットワーキング及び共同作業の、普及した効率的な方法になった。これらはブランド認知度の向上、顧客分析の改良、及び新しい販売経路をビジネスにもたらす。さらに、かつては互いに孤立的に暮らし、仕事をしてきたジャーナリスト、科学者、事業主及び一般市民が、ますます相互に結び付くようになってきている。ソーシャルメディアは、従来は交流範囲を超えると考えられていたような結び付きを即座に可能にする[15]。

しかし、AIは消費者に関する洞察を得るために採用されているだけでなく、既に人々の日常作業の一部として機能している。インターネット検索エンジンは既にAIを使用している。GoogleとBaiduは、いずれも検索クエリの精度を向上させる高性能アルゴリズムを開発した[47]。その他の適用は無数の分野で見られる。詐欺は、銀行口座を保護する機械学習アルゴリズムによって検出することができる[48]。Eメールのアカウントは、スパムを自動的にフィルタリングするアルゴリズムによって、よりクリーンに保たれている[49]。Facebookは顔認識を使ってユーザーを新しい画像と比較する[50]。Pinterestは画像内の特定オブジェクトを自動的に識別し、特定のカテゴリにそれらを割り当てることを可能にする[51]。TwitterとInstagramは、ユーザーの感情を分析するエンジンを開発した[52]。Snapchatは顔の動きを追跡し、動的な重ね合わせを可能にする[53]。その他、日々の生活の多くの例を挙げることができる。

これらのケースでのAIと人間の相互作用は受動的であるが、AIをより事前対応的にし、人と相互作用させる取組みも行われている。Siri, Alexa, Google Now又はCortanaは、自然言語処理(NLP)を使って、あらゆる種類の質問に答えるパーソナルアシスタントのようにふるまうことが可能である[54]。こうした性質は将来、確実により開発がなされるようになる。

第4節

人工知能の内部

この節では、AI が実際にどのように機能するのか、特に、現在どのような種類の機械学習メカニズムが一般に使われているのかについての洞察をさらに提供する。機械学習のアルゴリズムが扱う基本的な学習の問題の幾つか、及びそうしたアルゴリズムが実行される主なアプリケーション領域について、より詳しく説明する。その後、この節の最後の部分で、最も広く使用されている AI アルゴリズムの幾つかをレビューする。

4.1 機械学習のカテゴリ

機械学習はデータから情報を抽出する。三つの主要なカテゴリ、すなわち、教師あり学習、教師なし学習、及び強化学習が、このプロセスがいかに実施されるかの説明になる。

4.1.1 教師あり学習

データセットが既知の入力と出力のペアを含む場合、それは教師あり学習と呼ばれる。教師あり学習はこれに基づいて、未知のデータセットの出力値を予測するための一連の訓練データを使用する。教師あり学習を使用して開発されるモデルのパフォーマンスは、より良い一般化及び新しいデータセットに対するより優れた予測力の達成のために採用される、訓練データセットのサイズと分散(データ選択)に依存している。アルゴリズムは、分類(図 4-1)又は回帰(図 4-2)のいずれかを実行できる。

分類は、入力変数を離散出力変数にマッピングする。典型的な用途は、Eメールの中の特定単語の出現に基づいた、スパムのフィルタリングである(例えば、"\$\$\$", "金をもうける", "手数料なし")。回帰は入力変数を連続変数にマッピングし、例えば1年に渡る気温曲線の近似を行う。

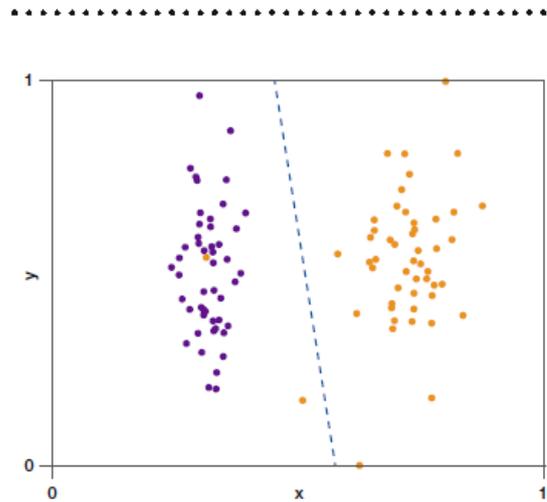


図 4-1 | 分類

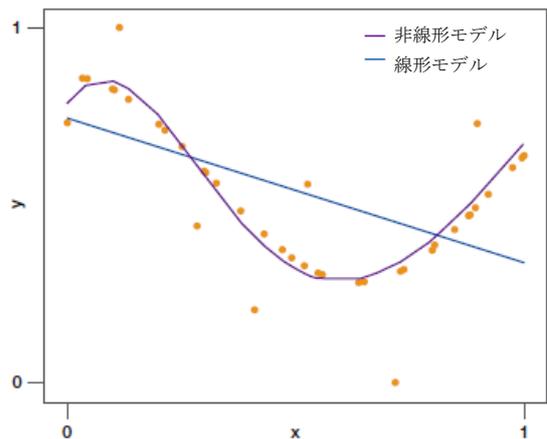


図 4-2 | 回帰

4.1.2 教師なし学習

教師なし学習モデルは、予め規定されている属性を定めることなく、データセットのインスタンスをグループ化することを学習する。このアルゴリズムは目標基準についての情報なしに、そのデータセットの基礎になっている構造を決定する。こうしたアルゴリズムは、一連の画像が与えられたときに、例えば様々な画像の中の対象物が同一ではないことを識別する。アルゴリズムは対象物について知ることなく、様々なカテゴリを形成させる。典型的なアプローチは、例えばクラスタリング(図 4-3)、及び次元を下げることを含んでいる。

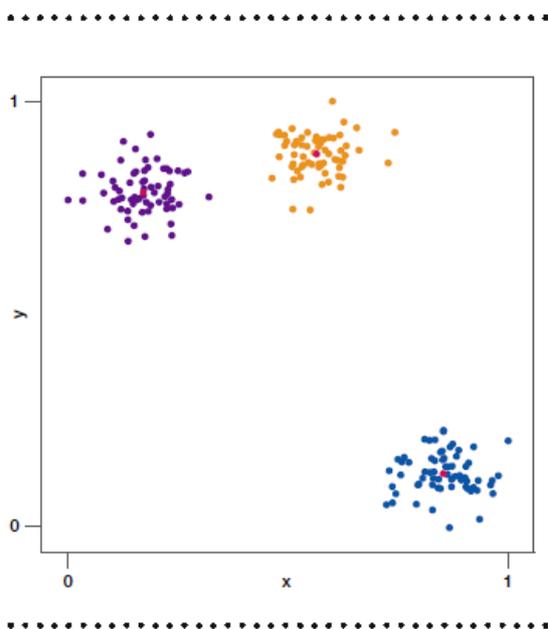


図 4-3 | クラスタリング

4.1.3 強化学習

強化学習は、エージェントの形態の AI システムが、その環境と、又は少なくともその環境の特定の対象物と相互作用するような学習方法を表す。エージェントはその環境で行動し、環境を観察することができる。エージェントは逆にその環境を通じて、通常は報酬の形でフィードバックを受け取る。この報酬

はポジティブな場合もネガティブな場合もある。エージェントの目標は、自らの行動を通じて、受け取ったポジティブなフィードバックを最大化するか、又はネガティブなフィードバックを最小化することである。これは行動の価値を概算する価値関数によってなされる。この関数は、とった行動とそれに対して受け取った価値によって改善されて行く。受け取るフィードバックとの関係で反復される行動によって、エージェントはより良い立場で、価値関数を通じて自らの行動の価値を概算することができる。フィードバックがどのように構成されているかに応じて、エージェントは特定の機能の実行について学べる。こうした学習エージェントのプロセスを図 4.4 に示している。

4.2 現在の機械学習システム

次の各節では、現在の一般的な AI システムの幾つかの例を説明する。

4.2.1 コンピュータビジョン

コンピュータビジョンは、物体又は人を識別するために人間の視覚システムを模擬するようなコンピュータの使用を意味する。これは通常、絵の中のパターンを認識し、そうしたパターンを画像の分類のために利用する機械学習アルゴリズムを使用している。コンピュータビジョンのタスクは、数値又は記号情報(例えば、決定)を生成するために、デジタル画像を取得、処理、分析及び理解する方法、及び現実世界から高次元のデータを抽出する方法を含んでいる。この文脈において、理解とは、他の思考プロセスとの仲介になり適切な行動を生み出すことができるような世界記述へ、視覚的イメージを変換することを意味する。画像データは、ビデオシーケンス、複数のカメラからのビュー、又は医療用スキャナからの多次元データなど、多くの形態をとり得る。CNN は一般にこの種のタスクに適用される。

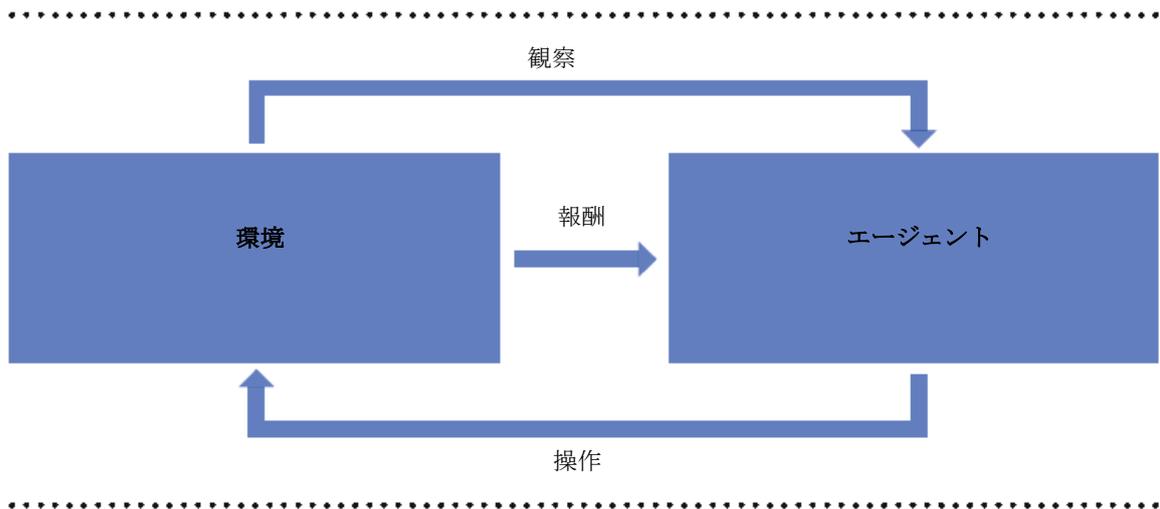


図 4-4 | 強化学習

4.2.2 異常検出

異常検出は、予想されるパターンからの逸脱を識別することが不可欠であるアプリケーションに使用される。これは、詐欺の見分け、ヘルスケアの監視、又はコンピュータシステムへの侵入の検出など、様々なシナリオにおいて見られるものである。機械学習は異常の迅速な検出を支援できる。

異常検出技術には三つの広いカテゴリが存在する。教師なし異常検出手法は、データセット中の大半のインスタンスが、データセットの残りの部分にほとんど適合しないように見えるインスタンスとは対照的に、正常であるという仮定の下で、ラベルのない試験データセット中の異常を検出する。教師あり異常検出手法は、既に“正常”及び“異常”とラベル付けされ、分類器の訓練を含んでいるようなデータセットを必要とする。最後に、半-教師あり異常検出手法は、与えられた正常な訓練データセットから正常な行動を表すモデルを構築し、その後、学習したモデルが生成する試験インスタンスの可能性を試験する。一般的に使用されるアルゴリズムは、 k 近傍法 (k -NN), SVM, ベイジアンネットワーク, 決定木, k 平均法を含むほか、長期短期記憶(LSTM)アプローチなどの ANN も含んでいる。

4.2.3 時系列分析

時系列分析は、一連の時系列データの中にパターンを見つけるための分析的アプローチである。その狙いは、ノイズによってあいまい化されている可能性があるデータ中の傾向を認識し、それを正式に記述することである。さらに、時系列分析は予測目的で、系列の将来の値の予測するために使用される。時系列分析に使用されるアルゴリズムには、隠れマルコフモデル(HMM), リカレントニューラルネットワーク(RNN), LSTMニューラルネットワーク, 及びSVMがある。

4.2.4 自然言語処理

NLP は、スマートかつ有用なやり方でコンピュータが人間の言語を分析し、理解し、意味を導き出す方法である。NLP を利用することにより、開発者は自動要約、翻訳、固有表現抽出、関係性抽出、感情分析、音声認識、及びトピックセグメンテーションなどのタスクを実行するための知識を体系化し、構造化することができる。

例えば音声認識では、コンピュータは話された言語を分析し、それを文章に翻訳する。これは、話された質問をコンピュータが受け取って答を探すよう

な言語処理システムの一部である場合がある。多くの在宅支援システムは、この種のユーザー入力に依存している。HMM 及び DNN などの手法がこのタスクにしばしば使用される。

機械翻訳において、ある言語から別の言語に文章を完全に自動翻訳する多くのプログラムは、自己学習アルゴリズムを使用している。翻訳における課題は、一般に、辞書で調べることができる各単語の意味が、特定の一つだけではなく文脈によって変化し得ることである。かつては統計及び規則に基づくモデルが機械翻訳によく使用されていたが、DNN の開発が近年急速に進歩し、しばしば優れた結果を与える。

4.2.5 推薦者システム

推薦者又は推薦システムは、ユーザーの好みに合ったその人向けのアイテムを予想する。推奨システムの普及は、多くの場合、デジタルのコンテンツ又はサービスの使用に基づくものであり、そうした場ではユーザーの好みを、与えられる評価に基づいてより簡単に特定することができる。協調フィルタリングがしばしば使用されているが、ナイーブベイズと k-NN アルゴリズムもこのタスクに広く使われている。

4.3 機械学習のためのアルゴリズム

現在の最も一般的な幾つかの機械学習システムの紹介は済んだので、次の 4.3.1~4.3.9 節で、これらのシステムを支援する主なアルゴリズムを説明する。

4.3.1 決定木

決定木は、分類と回帰、いずれのタスクにも適用できる。これらは通常、教師あり学習アルゴリズムの一形態として分類される。決定木は訓練データを使用して、決定ルールとその結果の概要を図式的に示す(図 4-5)。分類木が分類的又は離散的な結果を出すのに対し、回帰木は連続値を予測する。決定木は解

釈が容易で正確度が高いため、非常に普及している機械学習技術である。広く使用されている決定木アルゴリズムは、“反復ダイコトマイザー 3 (ID3)”, その後継である C4.5, 及び分類回帰木(CART)を含んでいる[55][56]。

これらのモデルは、動作モードが異なっている。予測を効率的に行うために、決定木の階層は、結果全体に対する属性の寄与に応じて定められる。結果に最も影響を与える特徴量は決定木の最上部に設定され、その後、より影響が小さい特徴量が続く。

一般に、決定木に関する一つの問題は過剰適合である。これは、モデルが訓練データセットのあらゆる詳細を抽出し、複雑すぎる決定木を構築した場合に起きる。その結果は、訓練データに対してはパフォーマンスが非常に優れているが、新しいデータセットに適合するための十分な一般化はできない。枝刈り法(ノードを作成又は決定木の深さを制限するのに必要な最少のサンプル数を設定)、又はアンサンブル法(異なったデータセットに対して訓練を行った複数の決定木を作り、それらを結合して単一のモデルにする)など、決定木の過剰適合を避けるための幾つかの解決法がある[58]。

4.3.2 サポートベクターマシン

SVM アルゴリズムは、教師あり機械学習の問題を扱う。これは分類と回帰のタスクに適用できる。このアルゴリズムの基本的なコンセプトは、異なるクラスを直線的に分割することであり、2次元で、直線的に分けられる二つのクラスの場合を図 4-6 に示している。このアルゴリズムは、データセットが与える諸クラス間の距離を最大化する。

これはハードマージン分類器と呼ばれる。最適な分類を行うために、このアルゴリズムは異なるクラス間の分離を最大化できるデータ点を使用する。諸クラスを分ける直線を定める選択されたデータ点はサポートベクトルと呼ばれ、アルゴリズムの名前はこれに由来している[57]。

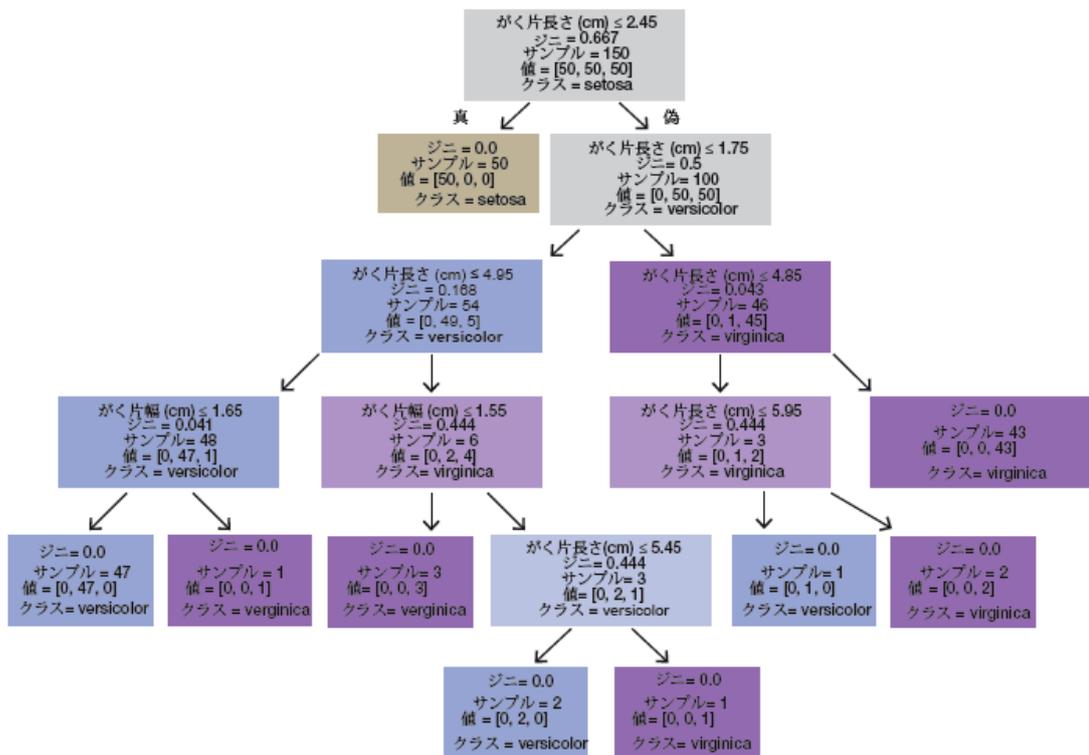


図 4-5 | 決定木

ハードマージン分類器の欠点の一つは、誤りが許されないために過剰適合が生じ得ることである。これは多くの場合、訓練データに関しては良いパフォーマンスであっても他のデータセットではパフォーマンスを低下させる。誤りが許されないことは、より高い次元において直線的な分離が可能でない場合にはアルゴリズムが解を見つけられないことを意味する。直線的でないデータセットの場合、ハードマージン分類器の代わりにソフトマージンが使用される。これは、分類の誤りを許容する変数を導入している。その後、変数はアルゴリズムのパフォーマンスを最適化するように調整される[4]。

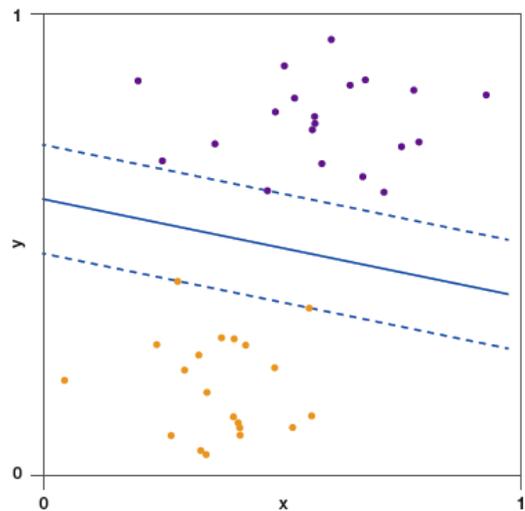


図 4-6 | 直線的に分割できるデータ(サポートベクターマシン)

4.3.3 ナイーブベイズ

ナイーブベイズ分類器は、ベイズの定理に基づく教師あり学習アルゴリズムのクラスである。データを分類するために、これら全てのアルゴリズムに共通している仮定がある。すなわち、分類されるデータの各素性は、クラスで与えられている他の全ての素性から独立している。ある素性の値の変化が別の素性の値に影響しない場合に、素性は独立している。

データセット中の訓練事例のクラスに基づいて、アルゴリズムは値を基にその特定のクラスに属する各素性の確率を計算する。新しいデータ点を分類する場合には、アルゴリズムは各素性のクラスの確率を別々に計算する。各クラスについてそれらの確率の積を計算し、最も確率が高いクラスを選択する。これが分類のための生成的アプローチである。

分類には区別的なアプローチの方が良いとしばしば論じられているが、ナイーブベイズ分類器は、比較可能な他の分類器よりもはるかに速く漸近誤差を達成できることが示されている[59]。

ベイズアルゴリズムは、テキスト検索又はスパム分類のような多くのタスクに適用される[60]。大きな長所の一つは、新しい素性機能に対するスケラビリティであり、特に、これは大きなデータセットに有用である[61]。

4.3.4 k近傍法

k-NN アルゴリズムは、一般に教師あり分類と回帰に使用されるが、教師なしクラスタリングにも適用できる。データは、新しいデータを分類すべき時までただメモリに保存されているので、このアルゴリズムは怠惰学習と呼ばれる。新しいデータは、保存されているデータ点に従って分類されるので、常にその時点の訓練データに依存している。

k-NN アルゴリズムの基本的なアイデアは、データを最小の距離で、k 個の最も近いデータ点とマッチさせることである(図 4-7)。分類のために考慮するデ

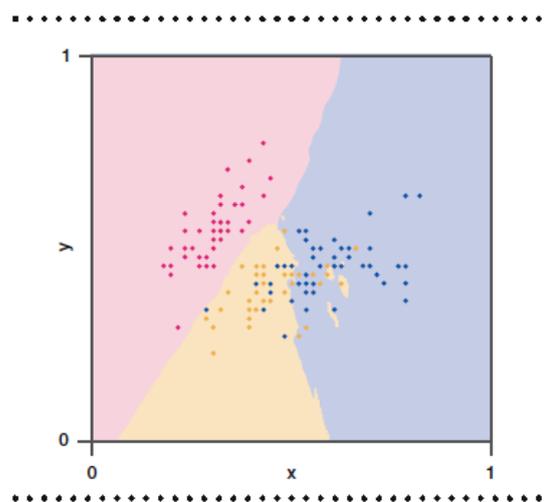


図 4-7 | k 近傍法

ータ点の数(k)の選択は、予測の質に大きく影響する。回帰アプリケーションの場合、予測は k-NN の値を使用して計算される。

k-NN アルゴリズムの予測の改善については多くの可能性がある。例えば、最も近い点の影響を大きくすることが、そうした点に重み付けをすることによって可能である。これは、より一貫性のある分類を可能にする。データの特性に応じて、異なった距離尺度がより良いパフォーマンスをもたらす。ほんの数例を挙げれば、ユークリッド距離は類似したデータタイプ(長さと幅)に適していることが示されており、マンハッタン距離は、体重、身長及び年齢のような異なったデータタイプを使用するときにより良く機能するように見える。

特徴とサンプルの選択は、k-NN アルゴリズムにとって非常に重要である。フィルタとラッパーは、最も重要な特徴選択手法の一つである。フィルタのアプローチは、アルゴリズム自体は実行せずに、無関係な特徴を除去することである。これは単に、データの構造を最もよく記述する次元を探す。一方、ラッパーは、アルゴリズムの結果を推定することによって、予測のパフォーマンスに最も関係する一連の特徴を明確化する[62]。

4.3.5 k 平均法

クラスタリングの問題では、ラベル付けされていないデータセットが与えられ、アルゴリズムはこれを自動的にコヒーレントなサブセットすなわちクラスタにグループ化する。[63]と[64]に示されているような k 平均法アルゴリズムは、この種のタスクに関して最も普及しているアルゴリズムの一つである。

k 平均法アルゴリズムは、クラスタのセントロイド(重心)と呼ばれるデータセット中の k 個のランダムな点を、ランダムに初期化することによって機能する。k の数は人が選択するか、又はある評価方法の使用によって定められる。その後、これは代入とセントロイドの位置変更という二つのステップで反復的に進行する[65]。クラスタの割り当てのステップでは、アルゴリズムは与えられたデータセット中のそれぞれの例を反復し、距離が最も近いことに基づいて各例を初期化されたセントロイドの一つに割り当てる。全ての例が、あるクラスタに割り当て

られるまで、これを各データ点に対して繰り返す。第二のステップでは、アルゴリズムは特定のクラスタに割り当てられている各データ点の平均距離を計算する。セントロイドはその後、計算された平均位置へ移動する。このステップを、k 個全てのクラスタについて繰り返す。このアルゴリズムを、クラスタのセントロイドがもはや動かなくなるまで繰り返す。これは、k 平均法アルゴリズムが k 個のクラスタへ収束したことを意味する(図 4-8)。

k 平均法アルゴリズムの結果は、全てのセントロイドが自らに割り当てられたデータ点をもつか、又は一部のセントロイドがそうでないかに応じて、k 個のクラスタ、又はそれ未満になり得る。アルゴリズムの結果は、セントロイドがランダムに初期化されるので変化する可能性がある。k 平均法アルゴリズムは、実際に適用する際、有用な結果を得るために、セントロイドの様々な初期化及び様々な k の数で複数回使用される。

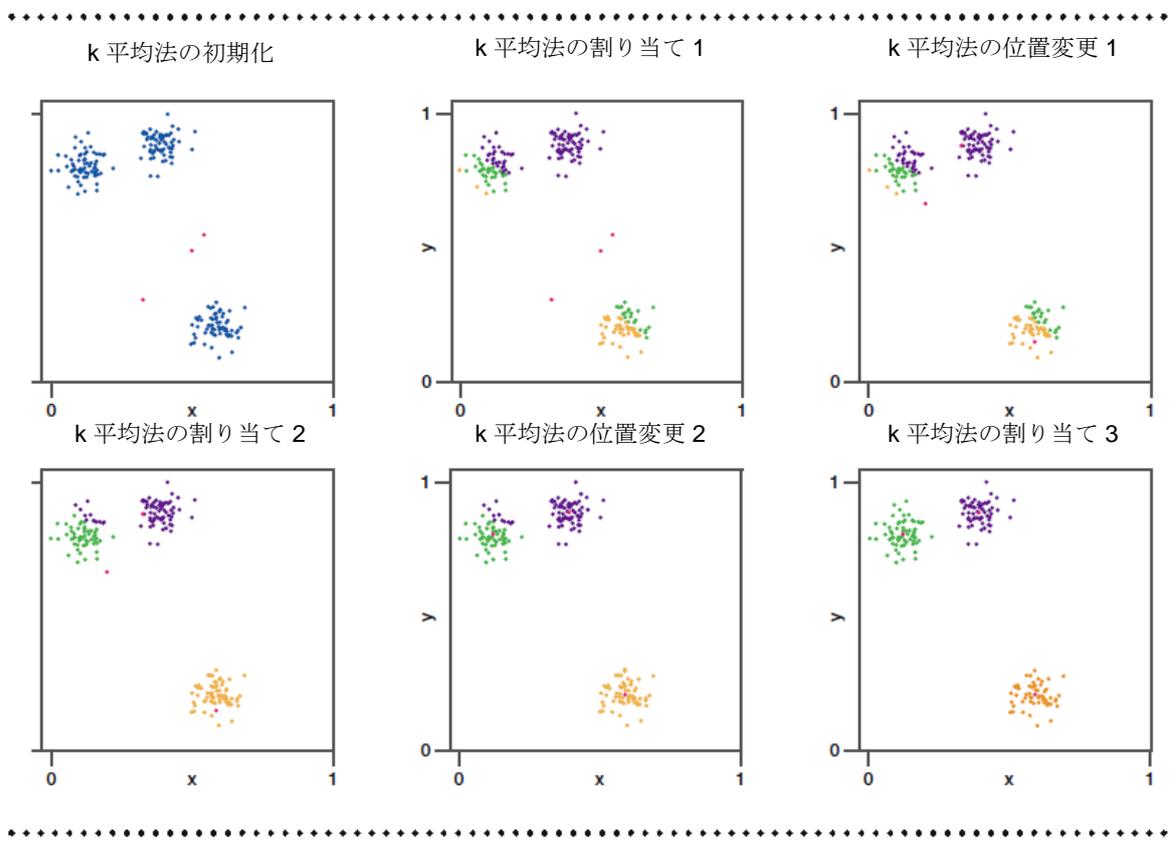


図 4-8 | k 平均法

4.3.6 隠れマルコフモデル

直線的シーケンスの確率モデルを作るための有用なアルゴリズムは HMM である。このアルゴリズムの基礎となっているコンセプトはマルコフ過程であり、これはシステムが一連の独特な状態にあることを、どんな時も記述可能であることを仮定している。間隔をもつ離散的な時間において、システムは状態に付随する一連の確率に従って状態間を変化する(図 4-9) [66]。

マルコフモデルの隠れた状態は、直接観測ができない確率過程であり、一連の観測を生成する別の一連の確率過程を通じてだけ間接的に観測することができる[66]。

HMM の適用分野は、DNA 及びタンパク質の分析におけるシーケンスモデリング[67]、情報検索システム[68]、及びオーディオシーケンシング[69]を含んでいる。

4.3.7 人工ニューラルネットワーク

ニューラルネットワーク(パーセプトロン)に基づくアルゴリズムは、AI の初期に開発された[3]。これは、教師あり学習と教師なし学習に適用できる。一般に、

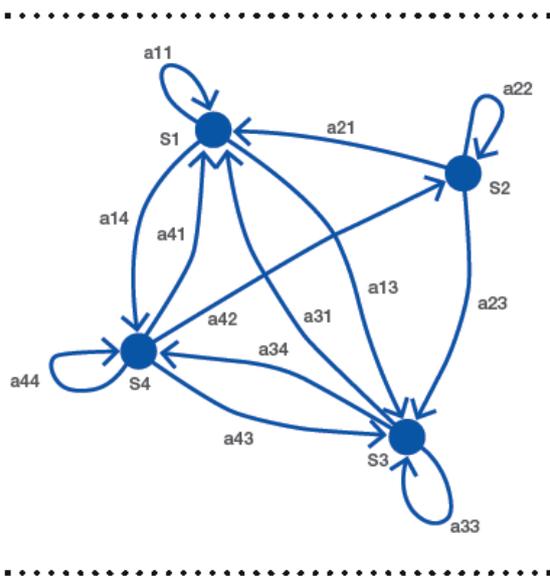


図 4-9 | 五つの状態の隠れマルコフモデル

ANN は人間の脳にヒントを得ているが、その機能をそのままコピーしているわけではない[70]。

ニューラルネットワークは様々な層から構成されており、各層は一つ前の層の全ての人工ニューロンに結合された人工ニューロンから構成されている(図 4-10)。入力層は入力データを表し、これは常に数値で構成されている。これは温度センサの出力のような構造化データ、及び画像のピクセルのような非構造化データを処理することができる。隠れ層のどのユニットが活性化されているかに応じて、出力層ユニットは予測を与える。画像認識の場合、これは例えば、画像中で識別された一匹の猿であるかもしれない。

人工ニューロンが各層のユニットである。それらは入力データを扱い、予測を可能にする。パーセプトロンの入力は、入力層からの元データ、又は DNN (複数の隠れ層をもつ) の場合は一つ前の隠れ層の人工ニューロンから変換された入力である。各入力は特定の重みによって調整される。その後、重み付き入力が処理され、セル本体で合計される。図 4-11 に示すように、偏り(固定された数)が調整変数として付加される。その後、セルの出力は、次の層への入力を表す活性化関数(この場合はステップ関数)の中で使用される。

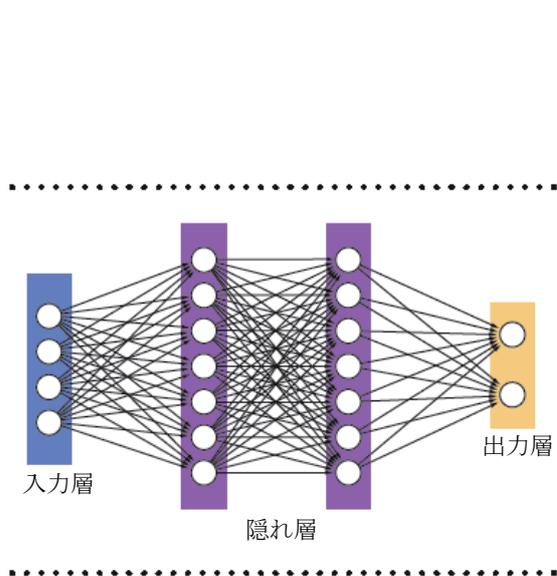


図 4-10 | 人工ニューラルネットワーク

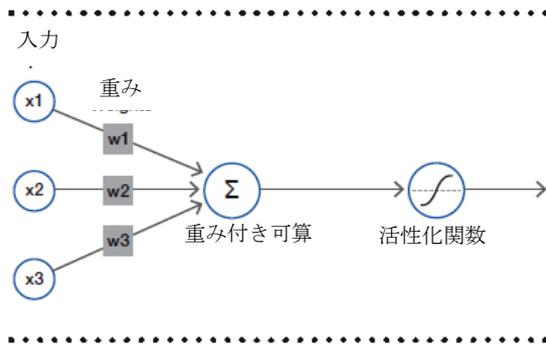


図 4-11 | 人工ニューラルネットワークのニューロン/パーセプトロン

ネットワークを訓練するために、重みはランダムに初期化される。次に最初の訓練データセットがニューラルネットワークへ入力される。その後、訓練の例の結果が、望ましい実際の結果と比較される。その後、逆伝播と呼ばれるアルゴリズムが重みの更新を行う。ただし、多くのニューロンと一緒に重なっている場合には、ネットワーク全体の最終出力の変化を制御することが非常に困難になる[71]。ニューラルネットワークを確実に機能させるための不可欠なステップは、連続的な活性化関数を利用して出力の変化を平滑化することである。ANNの主な長所の一つはその普遍性である。

4.3.8 畳み込みニューラルネットワーク

CNNは通常のANNと多くの類似点をもっている。これらは同じように、学習プロセスで調整がなされる重みと偏りをもつニューロンから構成されている。ネットワーク全体はそれでも単一の微分可能なスコア関数を表しており、最後の全結合層にはコスト関数が付加されている。しかし、通常のフィードフォワードニューラルネットワークとは対照的に、CNNは入力が画像であるという明示的な仮定で機能する。これにより、CNNはネットワークのアーキテクチャに特定の特性をエンコードすることが可能になる。これはフォワード関数の実行をより効率的にし、パラメータの数を大幅に減少させる[72]。

CNNは通常、畳み込み層、プーリング層、及び全結合層の三つの種類の層から構成されている[73]。畳

み込み層は入力ニューロンの空間的な関係を利用する。畳み込みニューロンの入力、入力層のニューロンの特定の領域から来る。この絞り込まれた受容野によって、畳み込みネットワークは従来のニューロンよりもはるかに焦点を絞った形で機能することができる。ニューロンの受容野は、フィルタのサイズと同じである。深さ軸に沿った結合の範囲は、入力ボリュームの深さに常に等しい。結合は空間的には局所的であるが、入力ボリュームの深さに沿っては常に全範囲である。

プーリング層は一般に、ネットワークアーキテクチャ内の連続した複数の畳み込み層の間に挿入される。その機能は、表現の空間的なサイズを小さくしてパラメータの量を減らすこと、さらには過剰適合を制御することである。プーリング層は、入力の深さの各スライスに対して独立に働き、各スライスを空間的にサイズ変更する。

全結合層のニューロンは、一つ前の層の全ての活性化されたニューロンと完全に結合している。これは、通常のフィードフォワードニューラルネットワークにあって一般的なことである。結果として、それらの活性化は、行列の乗算、及びそれに続く偏りのオフセットによって計算することができる。ほとんどのCNNアーキテクチャは、様々な数の畳み込み層、プーリング層、及び全結合層から構成されている。

4.3.9 リカレントニューラルネットワーク

RNNは特別な種類のANNである。それらは、教師あり学習と教師なし学習だけでなく、強化学習にも適用できる。ANNは、以前のデータからの独立性を仮定して現在の入力データを考慮するが、RNNは以前のデータを考慮できる。ANNのニューロンは前の層からの入力だけをもつが、RNNのニューロンは、その前の出力がループになっているためそれらへの依存性をもっている(図4-12)。このことは、この種のアルゴリズムが、例えば単語の文脈や時間的な側面など、シーケンス予測の問題を含むことを可能にする[74]。

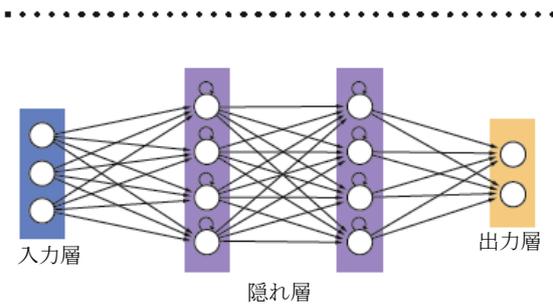


図 4-12 | リカレントニューラルネットワーク

これは、訓練段階で入力データの順序が重要な役割を果たすことも意味する。実際、RNN は入力を受け取り、現在及びその前の入力を使用してそれらの状態を計算する。保存されている以前の状態の全てが処理され、出力が計算されるまでこれが繰り返される。訓練の際に、得られた結果は実際の正しい結果と比較される。その後、ネットワークの重みを更新することができる。一般的な ANN との唯一の違いは、逆伝播は以前保存された全ての時間ステップを考慮に入れなければならないという事実である。逆伝播が時間と共にどのように機能するかの概要は、[75]に記載されている。

実際には、以前の全てのステップを共通の RNN に保存することは可能でないため、情報は時間が経つと共に失われる[76]。この問題に対処するために、双方向 RNN、LSTM、及びゲートリカレントユニット (GRU) のような様々なアーキテクチャが開発されて

きた。双方向 RNN は、以前の要素だけでなく将来の要素も考慮する。したがって、それらは順方向ループ用と逆方向ループ用の、二つのニューロンを使用する。その後、それらは次の順方向ニューロンに接続され、その逆も行われる(図 4-13) [77]。

LSTM ネットワークは、情報を保存できる、いわゆるゲートドセルを含んでいる。予測中に、セルはどの情報を保存、使用、又は忘却するかを決定する。入力及び出力ゲートは、訓練で得た重みに従って、情報を通過させるか又は遮断する。現在の入力、その前の状態、及びセルのメモリを組み合わせることにより、このアーキテクチャはデータセットの長期的な依存関係を明らかにできる[79]。

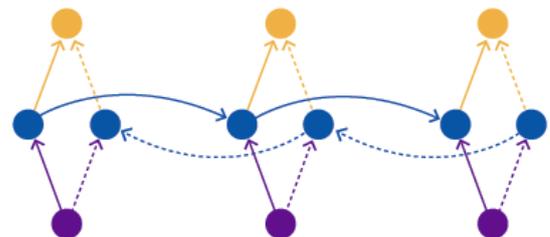


図 4-13 | 双方向リカレントニューラルネットワークのアーキテクチャ

第5節

人工知能の配備

チップ、データセット及びプラットフォームが成熟すると共に、AIの技術的及び産業的基盤は、より広く現実生活のアプリケーションを支援するために徐々に強化されてきた。画像認識とNLPのようなアルゴリズムの開発での技術的なブレークスルーは、AIがもたらす使用事例の数を増やし続けるであろう。マッキンゼーは、2025年までに世界のAIアプリケーション市場が総額1,260億米ドルに達すると予想しており、AIは多くのスマート産業の発展のためのブレークスルー点になるであろう[78]。

多くの伝統的産業の変化の中核的な推進力の一つとして、AIは意味認識、音声認識、顔認識等々の進歩によって築かれた、新しい技術と製品を生み出した。特定の産業のニーズの重要性を示すために、様々なAI機能に関する産業ニーズの概要をヒートマップ(色分け表)を使用して図5-1にまとめている。値のスケールは0~100であり、0はその機能が必要でないことを、100はその機能がその産業部門にとって極めて重要であることを意味する。この表は、様々な産業界における代表的AIシステムのマッピングを示している。

	コンピュータビジョン	異常検出	時系列分析	自然言語処理	推薦者システム
スマートホーム	80	50	30	100	75
スマートエネルギー	10	100	100	10	50
自動車	100	75	25	80	50
スマート製造	75	100	100	10	25
輸送/物流管理	100	75	75	10	50
金融	50	100	100	20	20
ヘルスケア	100	100	60	20	100
スマート農業	80	75	100	10	90

図 5-1 | 様々な産業界での一般的な AI システムの相互関係

5.1 スマートホームにおける人工知能

AI 技術の発展と共に、スマートホーム分野は、完全なスマートホーム・エコシステムを作るために家庭の各所に配置された個々デバイスを結合させる制御センターのコンセプトに向かって、少しずつ進化してきた。スマートホーム・エコシステムは IoT 上に構築され、ハードウェア(例えば、スマート機器、セキュリティ制御機器、家具)、ソフトウェアシステム、及びクラウドに基づくプラットフォームから構成されている。このシステムは、ユーザーのニーズを積極的に理解するために、音声認識、視覚認識、ディープラーニングのドメインモデル、ユーザーのポートレート、及びその他の技術的手段を統合している。

スマートホームは、デバイスの相互運用性とデバイスの自己学習を達成することを狙いとしており、ユーザーの挙動データを収集、分析することによって、家をより安全、快適かつ高エネルギー効率にするための、個人に合わせたサービスを提供できる。同時に、こうしたシステムは家電製品の効率も向上させ、エネルギーと天然資源の消費を減らし、より持続可能で健康的なホームスタイルを生み出すことができる。スマートホーム産業は、現在の家電市場の発展を促進し、AI の継続的な開発と産業化に貢献することもできる。

大手の家電メーカーは現在、スマートホーム製品を積極的に開発している。そうした成熟したアプリケーションは、スマート冷蔵庫、スマートエアコン、スマート洗濯機、スマート給湯器、スマートキッチン家電、スマートスピーカー、及び“全てが結合されている”というコンセプトを反映したその他の多くのスマート家電を含んでいる。企業は、様々な家電機器を接続して相互に制御することができる製品を開発し、予測及び分析作業のために多量のデータを収集している。インターネットをベースにした機能は一般に、消費者に広く受容されている。

スマートホームと AI の分野は密接に統合され、発展を続けている。機械学習、パターン認識及び IoT 技術の最近の進歩は、双方向性をより高いレベルへ引き上げ、住宅用デバイスをよりインテリジェントで

ユーザーフレンドリーなものにしている。製品は、携帯電話を中心としたものから、新しい革新的な人間-機械相互作用モードに焦点を当てたものへ、徐々に進化している。将来を展望すれば、AI 技術はスマートホームを受動的なインテリジェンスから能動的なインテリジェンスに移行させることができ、意思決定の一部では人間に置き換わりさえするかもしれない。産業界のステークホルダーのほとんどは、AI がスマートホーム市場に、より広範でエキサイティングな機会を開くことを予測している。

次の 5.1.1~5.1.3 節では、スマートテレビ制御システム、浴室セルフサービスシステム、及びインテリジェント食品識別システムという三つのスマートホームのシナリオについて説明する。

5.1.1 スマートテレビ制御システム

スマートテレビ制御は、インテリジェント、パーソナル、及び/又は省エネルギーのテレビサービスを提供することを目的としている。画像認識技術が向上したため、顔認識技術は著しく成熟している。指紋に比べて顔認識はより正確度が高く、アルゴリズムは光及びノイズのような環境要因の影響を、より受けにくい。スマートテレビ制御システムは、内蔵カメラによって家族の顔画像を集めて、パーソナルなサービスを提供することができる。例えば、システムは誰がテレビの前に座っているかに応じて、好みのテレビチャンネルに切り替えることができる。子供がテレビを見ているときに、ペアレンタルコントロールを自動的に作動させることもできる。

システムは、その時点のシーンに応じて最適な視聴距離を計算することができる。視聴者に最適な視聴体験を与えるように、音量、画面の明るさ及び彩度を自動的に調整できる。テレビは、観ている人が眠ったかどうかを検出し、眠った場合は音量、明るさ及び彩度を下げることができる。誰も観ていない場合には、システムはその点を尋ねることができ、応答がなければ自動的に電源を切ることができる。

5.1.2 浴室のセルフサービスシステム

浴室は、比較的プライベート、インタラクティブで、頻繁に使われる家庭の場所である。入浴では、人々は給湯器の温度を制御する必要があるかもしれない。しかし、水蒸気があるので機器に直接触れることはできない。そのため非接触制御機能が価値をもたらし得る。同時に、人々は入浴の際、散髪など、他に必要とすることがあるかもしれない。こうした状況においてスマートホームサービスは現代の家族生活を豊かにし、様々な地域及び文化に渡る多様な家庭のニーズに対処し得る。

一例として、中国市場向けのインテリジェント・セルフサービス浴室システムは、多方向の音声対話によって、便利な対話機能と豊かな情報をユーザーに提供する。入浴中、ユーザーは自らの声を使って、鏡の角度、シャワーの強さと角度、及び温度を便利に制御できる。同時にユーザーは音声制御により、タブレットのような他の家電製品から、美容、音楽及びフィットネス等々に関するコンテンツの推奨を受けられる。

5.1.3 インテリジェント食品識別システム

スマート食材認識システムを図 5-2 に示す。これは冷蔵庫内の食材の識別を行うために高解像度カメラを備えている。食材のそうした視覚的識別によって、システムはワンクリックでのショッピング、直ちに購入可能なショッピング、合理的な食事の計画、及び食材の早期リマインダーを支援する。これは、便利な操作、リアルタイム記録、健康的な食事の助言、及び家計の支出最適化など、幾つかの恩恵をユーザーに提供できる。

冷蔵庫とのリアルタイムでのやりとりを、携帯電話を通じて行うことができる。ユーザーは遠隔的に冷蔵庫の現在の中身を調べ、任意の場所で前もって食品の購入を計画できる。その他の恩恵は、食品の消費期限のリマインダー、冷蔵庫スペースの合理的な使用、あるいは食材をスーパーへ行かなくてもワンクリックで買えることを含んでいる。同時に、日々の食品消費量をクラウドデータベースに記録し、これをビッグデータの分析と組み合わせることで、食事傾向の評価と、バランスの取れたより健康的な食事の助言が行える。システムは、食事傾向、季節の食品、及び食事制限に基づいて、個人レシピを推奨することもできる。

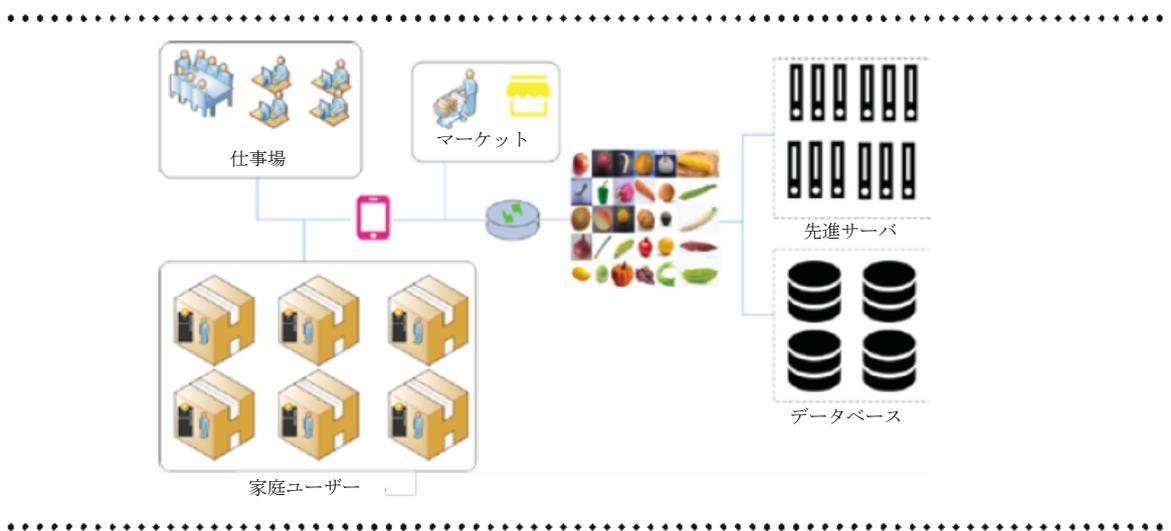


図 5-2 | インテリジェント食品識別システム

5.1.4 スマートホームの課題

スマートホームは、IoTの成熟度の高まりを基盤にし、ホームオートメーション分野にAIを付加するものであるが、リアルタイム学習、検出確率、組み込んだ場合の処理能力など、技術的な制限が存在している。その他の課題としては、マルチユーザーサポートの付加、及びデータの入手可能性の問題があり、これは機械学習のために信頼性の高い十分なデータを集める上で困難をもたらす。機械学習は通常、データの中の最も共通した内容を真実として扱い、一方、統計的にまれな内容は除外する。この認知方法は、人間のそれとはまだ異なっており、認知の偏りと誤解の発生につながり得る。

特定の問題を解決するために必要であるという視点から、データ総量中の信頼、信頼できる現実的データの割合が、機械学習で必要とされる割合に達しないことがある。これは特に、データをほとんどセンサ又はランプのような小さなIoTデバイスから得ているスマートホーム環境に当てはまることである。スマートホームのアルゴリズムは、様々なデータの流にに適応できる必要がある。例えば、ユーザーがテレビの位置をリビングルームから寝室に変更するときのようなデータの流の変化に、ロバストに対応できることが望ましい。それゆえ、高度なロバストネスが必要とされる。

個人化(パーソナライゼーション)、プライバシー及びセキュリティなどの問題も重要である。個人化とプライバシーは、ユーザーのプロフィール、サービスのカスタマイズ、及び日常生活でのデバイス設定に基づく。したがって、セキュリティ面で対処する必要がある。これは、スマート端末は不正侵入され、コントロールされてしまっている場合があり、個人情報漏れ可能性があるためである。さらに、AI対応のシステムも、人に対して害をもたらす得る。

浴室のような非常にプライベートな場所にスマートデバイスを設置する場合には、情報が秘密を保って処理されることの保証を伴っていないなければならない。そのデータから付加的情報が得られて人々のプライバシーを侵害することがあってはならない。そうし

た人々は、エネルギー消費を最適化するためにエネルギーデータの収集に同意するかもしれないが、そうした情報に基づく自らの日常作業の分析を、プライバシー侵害と見るかもしれない。

5.2 スマート製造における人工知能

スマート製造は基本的に、ICTと先進的な製造技術を統合したものである。設計、生産、管理及びサービスといった、製造に関わる全ての活動が影響を受ける可能性がある。

スマート工場は、未来の生産ラインの再定義に寄与するような高度に統合されたインテリジェントプラットフォームを作るために、サプライチェーン、設計チーム、生産ライン及び品質管理からのデータがリンクされた、ネットワーク化された工場である。多様な生産分野に適合できるフレキシビリティへのニーズが高まっており、製造部門は、生じつつあるこれらの課題に対処するために、自動化、機械学習、及びAIのその他の分野に頼る必要がある。

システムは機械学習を通じて経験から学ぶ能力をもち、結果としてシステムは常に向上している。これは、プラントの有効性から、最適な供給業者の選択、需要に照らした価格設定の判断に至るまで、全てを管理するための予測的な洞察を与えることにより、製造をより速く、よりフレキシブルで、特にスケラブルにすることを可能とする[87]。

製造業におけるAIのもう一つの恩恵は経済成長の支援である。そこでAIは、オンデマンド生産を実現し、運転効率を高め、製品サイクルを短くし、生産能力を高め、ダウンタイムを減らし、最終的にはコストを削減するために、労働要件と機械のスケジュールを含む資本効率の管理に使用される。

製造部門のAIの主なアプリケーションの幾つかを、次の5.2.1～5.2.4節で説明する。

5.2.1 予測保守

伝統的な製造ラインは、生産機器の障害の警報が生じたとき、既に品質に問題のある製品を多数生産してしまっていることがあり、これは企業全体に損失をもたらす。リアルタイムで収集される機器の運転データにより、予測保守は機械学習アルゴリズムを通して障害信号を識別でき、欠陥機器の早期検出と保守を行うことができる。最終的に、これは保守の時間と機器コストを削減し、機器の使用を改善し、機器故障による損失を回避させる[88][89]。障害の予測と障害位置の特定及び診断は、予測保守のための二つの主要メカニズムである。

• 障害の予測

工場のデバイス又はネットワークの主要業績評価指標(KPI)は、通常、徐々に悪化する傾向を示す。通常は、ハードウェア又はサービスの障害が生じる前に、不安定な又は劣化した運転状態が現れる(図 5-3)。障害の後の受動的な処理は、サービスエクスペリエンスに影響するだけでなく、トラブルシューティングにも長い時間がかかる。

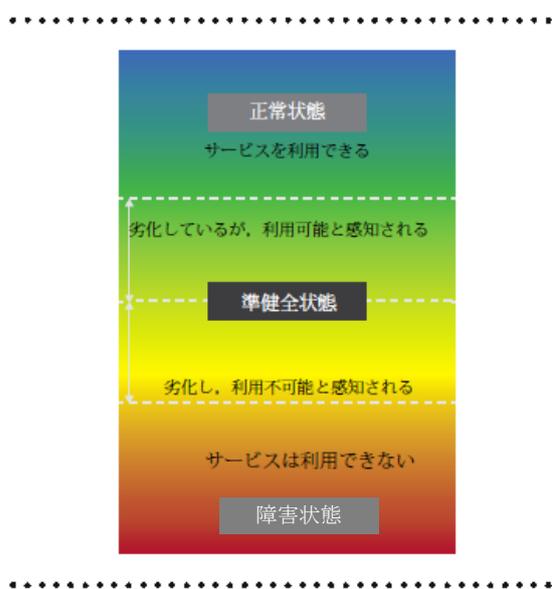


図 5-3 | デバイス/ネットワークの健全性

デバイス又はネットワークの KPI 予測を通じて、サービスの中断、ネットワークリソースの非効率性及び悪化を防ぐことができる。図 5-4 に示すように、デバイス又はネットワークの障害の予測には二つの主な方法がある：

- ケース 1：黒い曲線は、“今”の時点よりも前の、そのデバイス又はネットワークの健全性状態を反映している。赤い曲線は、黒い曲線の履歴データから得られる傾向である。この傾向が、あるしきい値を下回ると警告が生じる。
- ケース 2：黒い曲線はケースと同じである。赤の曲線は、予測アルゴリズムと黒の曲線からの履歴データに基づいた予測曲線である。逸脱が、あるしきい値を超えると、異常を報告するために警告が発生する。

• 障害位置の特定及び診断

工場内のデバイスに欠陥がある、又はそれが不適切に運転されている場合には、障害が拡大してネットワーク障害又はサービス障害を発生させることがある。障害の位置を素早く特定することは、障害回復時間を短縮し、損失を減らすことに役立つ。

工場でデバイス/ネットワーク/サービスの障害が生じた後、ログ、警報、KPI、構成及びトポロジーなどのネットワークのパフォーマンス及び状態データが収集される。その後、障害デバイスの位置を素早く特定し、障害の根本原因を特定できるように、相関分析を実施してどのメトリック又はパラメータが異常かを判断する(図 5-5)。

障害位置の特定及び診断については、二つの方向の相関分析が存在する：

- 水平方向では、ネットワークサービストポロジーに従って、関係するライン分析を実施するためにパス上の全てのデバイスのメトリックをまとめる必要がある。
- 垂直方向では、関係するライン分析を実施するために、デバイスの物理層メトリック、ネットワーク層メトリック、及びアプリケーション層メトリックをまとめる必要がある。

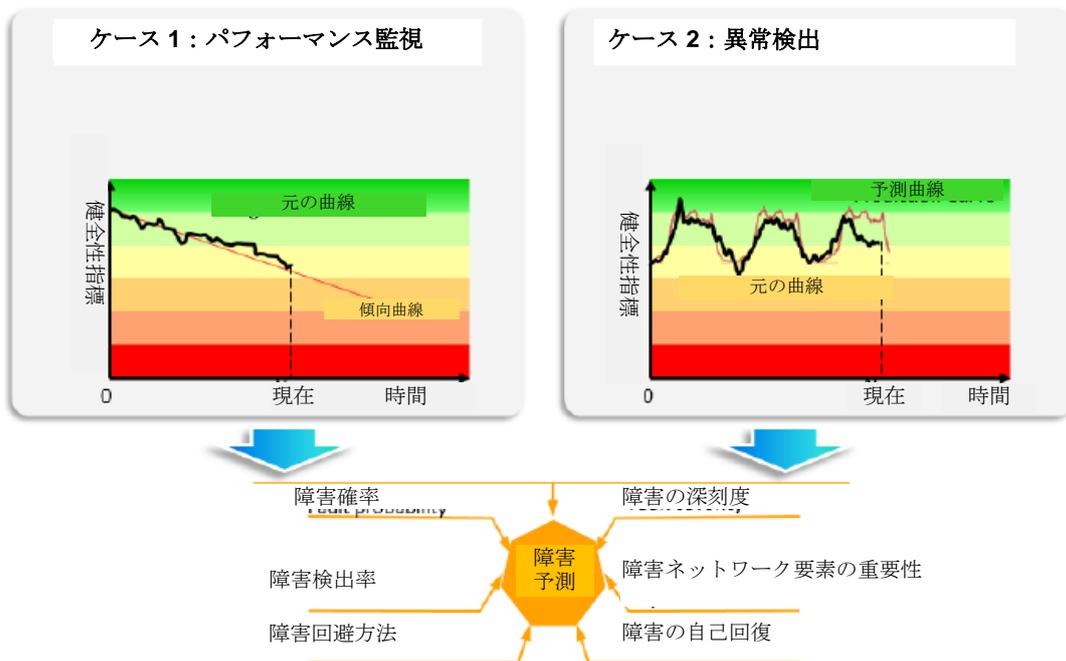


図 5-4 | デバイス/ネットワーク障害予測の二つのケース

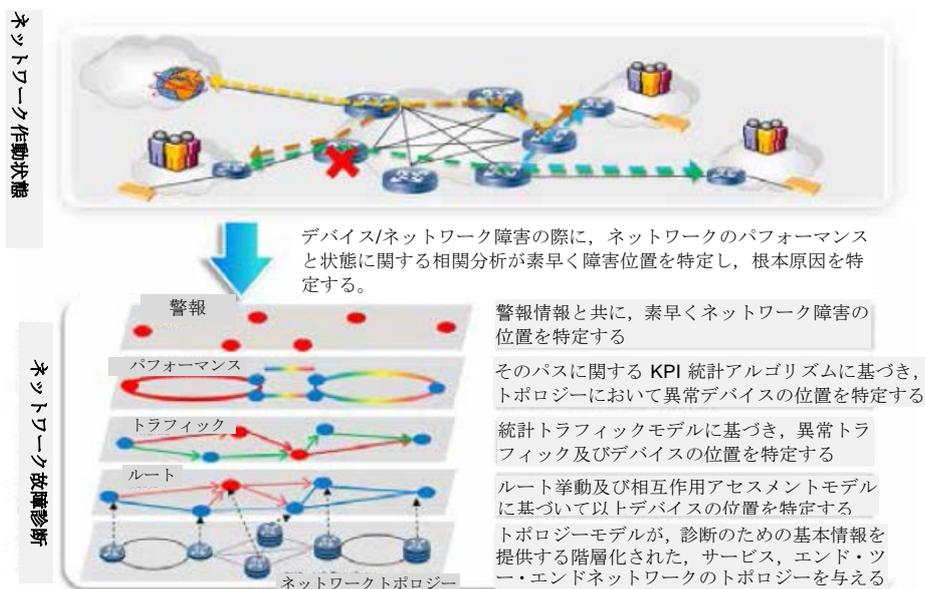


図 5-5 | 相関分析に基づく障害位置の特定及び診断

させ、検出漏れ又は誤った検出に起因する損失を回避する。自動化された品質管理により、チップ製造などのアプリケーションにおける欠陥率を大きく減らすことができる。同時に、何が製品欠陥の原因であるかを分析して、製品の棄却率を下げ、品質管理コストを下げるために製品の設計と製造を最適化することができる。

自動画像認識(AIR)は、集積回路の製造最終工程からの説明的事例である。AIR システムは、初めは X 線装置を使用した自動画像検出のニーズに対処するために開発された。コンピュータソフトウェアと画像処理ハードウェア技術を組み合わせることで、視覚的検出の精度と信頼性が向上し、最適化され、それにより人の目への過度の依存を避けることができる。

半導体デバイスの内部構造を視覚化するために X 線装置が使用される。一般に、一組の製品は 20~50 リールを備えており、5~10 枚の画像を撮影すると 1 リールで全てのデバイスをカバーできる。画像は最終的に、インターネットに接続されたストレージサーバに保存される。画像パターンの例を図 5-7 に示している。

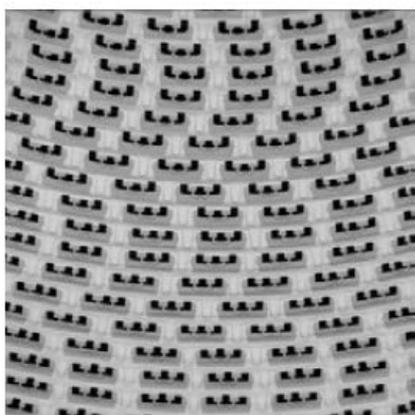


図 5-7 | パターン撮影の元画像

AIR サーバは、中核となる画像識別アルゴリズムとオペレータコール、事前画像評価(これで適格な画像にのみ、その後の自動識別が行われる)、画像撮影のカットオフ時間の追跡、製品情報の照合など、幾つかの機能から構成されている。無効な画像は、フォローアップとして人によってなされる二次的な目視検出フィルタリングのために、自動的に取り出されて再配置される。

サーバの自動識別モジュールが元画像をフィルタリングする。システムがフィルタリングした、障害が疑われる少数の装置を人の目で再チェックして、実際に故障している装置を見つける。図 5-8 に、故障が疑われる装置のリストと、再配置後のその位置の例を示している。オペレータは実際に故障している装置を特定し、リールの物体の検出に便利のようにインタフェース上でそれらをマークする。

クライアントは、追加で手直しを行うためのユーザーインタフェースを備えており、これは視覚的検出による欠陥製品の再チェックを要求するために使用される。画像の質の可否チェックに関し、各 X 線装置が生成する画像の質(明るさ、グレースケール、コントラスト)をエンドユーザーがアセスメントするための機能をクライアントは備えている。システムは、自動識別の要求事項を満たすために撮影パラメータの調整を行うことをユーザーに促す。



図 5-8 | 視覚的検出

5.2.4 スマート製造における課題

スマート製造には、特に AI 関連の技術を適用するとき、幾つかの課題がある。大きな関心事の一つは、“ヒューマンインコマンド(人間が指揮する)”という言葉で要約することができる。悪用又は誤用を避けるため、スマート製造は AI にヒューマンインコマンド・アプローチを採用することが望ましい。これは、機械は機械としてのあり方に留まり、人間が常に機械に対する制御を保持することを要求するものである。

協力ロボットに対して AI を使用する際の最大の課題は、データ取得とアルゴリズムのロバストネスである。アルゴリズムが常に意図した通り確実に動作するようにする必要がある。特に、リスクがある状態で人間と協力して仕事をする場合には、システム全体が常に予測可能な形で動作しなければならない。とりわけ、ニューラルネットワークを使用するとき、これは未だ大きな課題である。これにより、特に入力に予期しない変化がある場合に予想外の結果が生じる可能性がある。この問題は、例えばニューラルネットワークが極小値へ下がる時に発生する。開発者はこうした極小値に基づくネットワーク決定を、多くの場合意図していない。

AI とロボティクスの発展は産業の競争力を高め、次にはそれが新たな仕事を創出し得る一方、人間が現在行っている仕事が次第にロボットに代替されるという重大なリスクも存在する。労働市場の動的変化に照らせば、教育と訓練(職業訓練、生涯学習及び訓練、職場での再訓練を含む)が、将来はいつそう重要になる。ロボットと自動化の利用の増加に対応するには、新しいスキルと力量が必要とされる。

スマート製造のその他の課題は、AI アルゴリズムを訓練するための質の高いデータを大量に利用できること、及びそうしたデータを意味のある情報とドメインモデルへ構造化する必要があることなどを含んでいる。

5.3 スマート輸送と自動車部門における人工知能

スマート輸送は非常に幅広い市場潜在力(例えば、自家用車、公共交通機関、駐車場、物流管理、緊急サービス)をもち、AI 技術を必要とするということが広く認識されている。特に大都市における車両数の激しい増加に伴って、交通管理と渋滞制御などのサービス、及び自動運転車の最近の急増が、いずれも AI 技術の大規模な使用を求めている(例えば、画像解析、ルートの最適化、対象物の識別)。図 5-9 に示すように、AI が提供する機能は、エンドデバイス(例えば、携帯電話)だけでなく、エッジ(例えば、車)及びクラウド(データセンター)でも採用されている。

5.3.1 自律運転

無人運転車としても知られている自動運転車は、自動運転を行うためにインテリジェントコンピュータシステムを使用する車両である。こうした車両は、インテリジェントな経路計画技術、コンピュータビジョン、及び全地球測位システム技術によって、車載コンピューティングシステムが安全かつ人間の介入なく動作することを可能にする。

自動運転車は急速に発展しており、業界の巨大企業及び新興企業は、短期間の内に自律走行車を計画又は発売する準備が既にできている。高速道路交通安全事業団(NHTSA)が、五つのレベルの車両の自動化を定義している[95]。

- 非自動化(レベル 0)。運転者が常に完全に制御を行う。
- 機能特有の自動化(レベル 1)。一つ又は複数の特定機能の自動化。
- 複合機能自動化(レベル 2)。共に作動するように設計された少なくとも二つの主な制御機能の自動化。
- 制限付き運転自動化(レベル 3)。運転者は、特定の条件下で、安全上重要な全ての機能の制御を完全に機械に委ねることができる。車両は、運転者による制御に戻る必要がある状態を監視する。

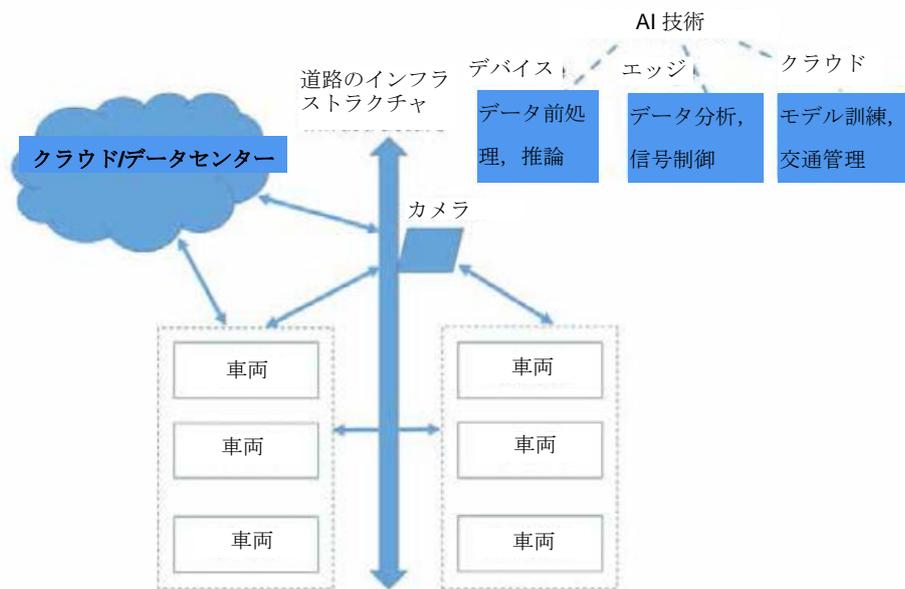


図 5-9 | スマート輸送における AI 技術のアプリケーション

- 完全運転自動化(レベル 4)。車両は、安全上重要な運転機能の全てを実行するように設計されている。

明らかに、自動化のレベルが高ければ高いほど、運転の責任を車両がより多く担う。これは、AI 技術が車両の内部とクラウド、両方で必要とされることを意味する。そうした技術は下記を含むが、これらに限定されるわけではない(図 5-10) :

- 車両を制御すること。これは、人間が介入しない始動、ブレーキ、旋回、及びその他の自動操縦機能を含んでいる(運転者の負担を除去する、又は人間が介入できないときに介入する)。
- 運転者の状態、車の状態、道路の状態、及び周囲環境(例えば、歩行者、動物、障害物)を識別すること。周囲の物体の識別と分析は、広範な機械学習機能を必要とする。さらに、運転の手助けをするには、分析は非常に低い待ち時間で行われる必

要がある。実際のところ、従来の汎用 CPU は、こうしたシナリオにおいて AI アルゴリズムで必要とされるコンピューティングパフォーマンスと効率を提供できない。モデルの訓練、推論、及び AI の作業負荷に合ったエミュレーションのために、専用のアクセラレータが(車内とクラウド、両方で)必要である。

- フリート管理、すなわち組織化された車両の一群の自律操縦を可能にすること。これは、物流管理と配達サービスにおいて広く使用できる。

5.3.2 交通管理

交通管理はスマート輸送において不可欠な役割を果たしている。世界中で人口と車両が増加するにつれ、都市部における大量の交通の管理は、政府、警察及び自動車会社にとって、困難ながら不可欠である作業になっている。エコノミスト誌の報告[94]によると、交通渋滞に起因する支出は、2013 年にフランス、ド

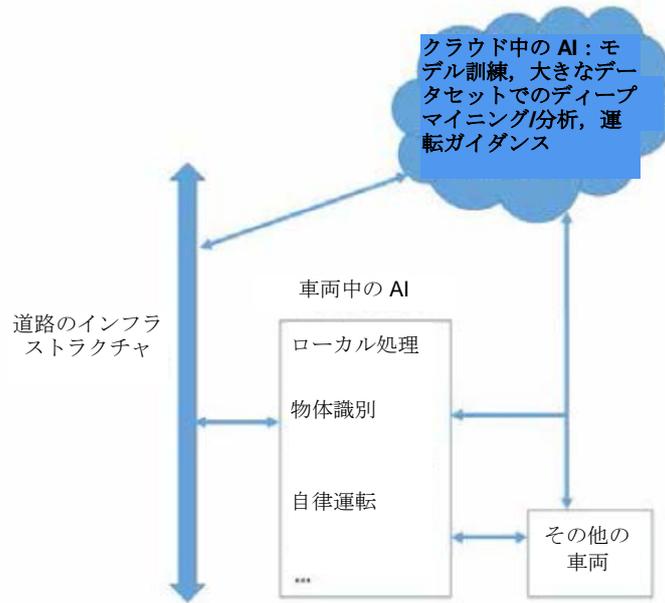


図 5-10 | AI 技術を用いた自律走行車の実現

イツ、英国及び米国で既に総計 2,000 億米ドル(GDP の 0.8%)に達している。したがって、交通の流れを最適化し、交通渋滞を減らし、車の排気を最小にすることは、生産性の向上、生活の質の向上、及び環境保護に大きく貢献する。これらの目標を達成するために、AI 技術のより広く、より良い活用が必要とされている。

アプリケーションの例は、下記を含んでいる：

- 交通の流れの分析。機械学習とデータマイニングを使用することにより、複数の道路での、又はより広い領域でのリアルタイムの交通データ(例えば、車両、歩行者、渋滞、事故)を、交通ルートの最適化、渋滞を避ける交通整理、及び排出の削減のために分析し、相互に検討することができる。

- 信号機の最適化。交通信号の切り替え(これは現在、リアルタイムの交通状態を考慮していない)を静的に決める代わりに、AI アルゴリズムを使用してリアルタイムの交通状況を分析、予測し、動的な交通信号の切り替えを行って車両及び歩行者の通過を最適化することができる。
- 規則と規制に対する違反の検査。これらの作業は従来、人間による作業を非常に多く含んでいる。画像と動画の処理に AI アルゴリズムを用いてさえ、リソースと計算能力の不足のためにパフォーマンスは今も限られている。しかし、AI の作業を加速するより強力な処理プラットフォームを使用すれば、動画と画像をはるかに大量に分析することができ、人件費の削減、精度の向上、及びパフォーマンスの向上につながる。

その他の例は、スマートパーキング、監視カメラ、又はスマート物流管理を含んでいる。

5.3.3 交通ロボット

輸送における AI のもう一つの開発は、道路の交差点に交通整理のために配置されるインテリジェント交通ロボットの使用である。こうしたロボットは AI 技術を使用して交通の状態を監視し、交通情報を取得し、問題の起きやすい道路交差点において交通に関する指示を出すことができる。ロボットは歩行者に対して、交通規制を守り、安全に対する意識を高めることを促し、手による指示、ペンライト、声による警告及び安全メッセージの発信によっている交通警察の作業負荷を減らすこともできる。さらに、ロボットは歩行者の違反を記録するために画像認識を使用し、それにより交差点での全体的な安全性を高めることができる。

5.3.4 スマート輸送における課題

安全、セキュリティ及びプライバシーは、スマート輸送に関するユーザーの大きな関心事である。各車両に搭載されたセンサと自動運転のための先進的な AI 技術の実装により、データ、及び人間の運命さえコンピュータに委ねられている。最近の自律走行車の事故はさらに懸念を生じさせている。そのため、安全、セキュリティ及びプライバシーの問題は、スマート輸送にとって最も重要な課題の一つである。

スマート輸送に関連する政策、法律、規制及び規格は、対処する必要があるもう一つの重要な課題である。規格は、例えば収集したデータの安全な伝送、保存及び処理を確保するために、スマート輸送技術の実装に関係することがある。規制又は法律は、例えば、自動車メーカー又はサービスプロバイダによるスマート輸送における AI の使用に関係することがある。

もう一つの課題は、下記の目的のためにクラウド、エッジ及びエンドデバイスの AI 作業負荷に適合させた、異種コンピューティングプラットフォームが利用可能であることである。すなわち、1) 自動車、歩行者及びインフラストラクチャからの増大しつつある量のデータを統合して、それらのデータを効率よく送信、保存及び分析することができる。2) AI の作

業負荷(例えば、モデルの訓練と推論)を迅速に効率よく処理でき、交通管理と自律運転の(半)リアルタイムの意思決定を支援することができる。プラットフォームのアーキテクチャは、一方では、AI の作業負荷とアルゴリズムの進化に追いつくように改善又は革新される必要がある。他方では、ユーザーとアプリケーションの要求の増大(例えば、待ち時間、データの量、接続性)を満たすために、パフォーマンスが絶えず改善される必要がある。さらに、プラットフォームは新しいコンポーネント又は機能を組み込むために、スケーラブルである必要もある。

結局のところ、フル AI スタックに対して急を要する必要性が生じるであろう。全てのサブシステム(すなわち、エンドデバイス、エッジ及びクラウド)において、画像分析及び顔認識のような機能が必要とされるかもしれない。例えば運転者は、自らの状態に関するデータの前処理のために、自分の携帯電話に AI 対応ソフトウェアをインストールするかもしれない。エッジコンポーネント(例えば、車載の)は、車両が生成するデータと環境から収集したデータをより多量に処理することができる。その間にも、クラウド/データセンターは、ルートの最適化、交通案内、又はその他の形態のマクロ管理などの分析を実行するために、全ての車両、交通の流れ及び環境からのデータ、並びに履歴データを大量に収集する。

5.4 スマートエネルギーにおける人工知能

エネルギー部門は、一次エネルギー(例えば、石油、ガス、石炭)と二次エネルギー(例えば、電気)に分類できる。前述のように、天然資源の枯渇と気候変動の影響は、さほど遠くない将来に、現実となりつつある。これらの問題を解決するために、世界の国々はエネルギー需給管理の最適化に焦点を当てた対策を実施している。

スマートエネルギーの目的は、エネルギー生成、輸送及び消費の歩留まり率を高めるだけでなく、効率のよいエネルギー管理を可能にすることである。これまでは、できる限りエネルギー損失を減少させるということに重点が置かれていた。IoT、ビッグデータ及び AI の成長を含む ICT の収斂は、エネルギー部

門で技術革新を成し遂げる推進力になりつつある。これは、生産段階と消費における需要と供給の推定を向上させ、様々な建物と機器の間でエネルギーをやりとりするサービスを可能にする。さらに、通信技術によって収集した情報を分析することで、省エネルギーの最適化はこれまでなかったレベルに達する可能性がある。IDCによれば、スマートエネルギー市場は2020年までに5,160億米ドルに達する可能性がある。

5.4.1 グリッドの管理と運用

スマートグリッドは、供給業者と消費者の間のエネルギーのやりとりを、ICTを使用して可能にすることにより、エネルギーの生成、輸送及び消費の効率を向上させるインテリジェントな電力グリッドシステムである。従来の電力グリッドでは、電力は大きな発電所から最終消費者へ一方的に供給されていたが、スマートグリッドでは双方向の電力の流れが生じる。言い換えれば、最終消費者は、電力が必要なときは外部電力の消費者であるが、例えば再生可能エネルギーの生成によってローカルに余剰電力があるときは、電力をグリッドに戻すことができる供給者でもある。

スマートグリッド内の電力機器のリアルタイムの監視、操作及び保守を支援するために、AI技術を使用することができる。AI技術は、機器の動作の安定性と発電効率を向上させるために、障害の初期段階において障害の診断機能と修復対策を提供することができる。機械学習は、効果的な予防保守を実施するために、様々な動作条件においてごく小さなパターン変化を識別することを可能にする。

機械学習の強力な予測機能を活かして、世界中の様々な公益事業と大企業が、自らのエネルギーインフラストラクチャの効率を最適化して、一般に10%~15%高めることに成功した。AIは、特に分散型再生可能エネルギー発電を扱う場合に、電力のピーク需要と供給を高い精度で予測することもできる。これは、消費者がエネルギー料金を大幅に減らす助けとなり得る。

5.4.2 消費者の関与とサービス

スマートグリッドは一般に、分散型の先進的計量インフラストラクチャ(AMI)と、エネルギーの生産者と消費者の双方向通信を支援するスマートメータを備えている。様々なAIアルゴリズムを使用して電力生成と消費者物価を適合させるために、ウェブ・ポータルを使用し、様々なエネルギー関連のデータを収集、表示したり、消費性向を明確化したりすることができる。一方で、先進的プラットフォームは産業界の大規模消費者が、その特定のビジネスニーズに合った、より洗練された電力取引を行うことを可能にする。

さらに、AMIは、電力使用パターンの遠隔分析によって、高齢者又は障害者をモニタリングするような新しいサービスの機会を提供する。シナリオは、例えば深夜に照明が消されない、又は朝に電力使用が増えない場合に、関係する人又は団体に注意を促すことを含んでいる。

5.4.3 統合スマートエネルギープラットフォーム

消費者の関与をいっそう高めることが必要とされるため、世界の多くの公益事業は、AI技術が可能とする広範なサービスを統合した、スマートエネルギープラットフォームを採用している。こうしたプラットフォームの一例は、韓国に拠点をもつ公益事業、KEPCOによるHUB-PoP(ハイパー接続されたユビキタスブリッジプラットフォームのプラットフォーム)である。図5-11に示すように、HUB-PoPは、電力グリッドの運用、管理支援、顧客管理、再生可能エネルギーを含む新しいエネルギービジネスなど、様々なサブシステムを接続する統合的なクラウドベースのプラットフォームを提供する。顧客に様々な革新的サービスを提供するために、プラットフォーム上でデータ分析とAIアルゴリズムを実行することができる。



図 5-11 | 電力管理のための AI ベースの HUB-PoP

こうしたサービスの一つは、複数の顧客の要求と並行して処理を行うことを可能にする自動チャットボット対話システムである。チャットボットは機械学習技術を使用し、継続的な改善のために、相談とオンラインの質問を通じて蓄積された知識を得る。

5.4.4 スマートエネルギーにおける課題

幾つかの使用事例で示されているように、エネルギー分野の様々なアプリケーションとサービスに AI 技術を適用することが可能である。しかし、エネルギーインフラストラクチャの規模と寿命の長さから、新しい機器と技術の導入は速度が遅く、しばしば複雑である。適切な投資収益率の達成は一般に、動きの速い他の産業よりもはるかに時間がかかる。多くの場合、公共部門と民間部門の効果的な協力が、そうした技術投資を妥当なものにする前提条件である。

さらに、あまり調整されずに実装される需要側リソースが出現すると、これはグリッドを不安定化させ、極端な場合は設計容量を超えるフローを生じさせる可能性がある。これが、中央グリッドの計画と設計を進める際に、AI が役割を果たすことができる部分である[93]。

入力データが分散していることが、エネルギーネットワークの分析における一つの課題である。こうしたデータは、現在、様々なプラットフォーム上で分散した形で入手される。入力データの間の相関を AI アルゴリズムで検出するには、まずそれらを収集し、標準化する必要がある。こうしたデータが統一的に集められていない場合にも、問題が生じ得る。例えば、時間不変なものとしてデータが収集されたかどうか、又はデータがどのようにスケーリングされているかを考慮する必要がある。多くの場合、データの取り扱い及び評価も、そうしたデータが個人のプライバシーの分析に使用できる場合には、重要な役

割を果たす。電力消費とパターン分析は、ある人物がいつ家にいるのか、及びその人物が何をしているのかを判断するために使用することができる。さらに、先進的な分析は、そのことがはっきり表明されたり望まれたりしていなくても、エネルギー使用挙動の予測に基づいて家の中に誰がいるかの判断にも使用できる。こうしたデータプライバシーの課題は、電力会社が適切なセキュリティ対策と政策を用いて対処する必要がある、大きな懸念である。

P60

第6節

人工知能の課題

前節で説明した諸産業で AI 技術を普及させるには、技術、倫理、信頼性及び規制に関連する幾つかの課題に対処する必要がある。例えば、輸送及び自動車部門で AI を実装する場合には、安全とセキュリティが最も重要な課題の一つである。スマート製造の場合には、安全と信頼性が大きな関心事である。

例えば、倫理的及び社会的影響、AI インフラストラクチャの計算能力と効率、データの利用可能性と質など、幾つかの課題は様々なアプリケーション分野で共有されている。これらの課題に対処することが、複数の産業に渡って AI 技術の普及を加速する助けになる。

6.1 社会的及び経済的課題

AI は社会と市場の両方に深く影響を与える可能性がある。AI は製造業だけでなく、法律の実務のような手続き的な仕事も含む幾つかの分野で自動化を可能にすることにより、仕事の破壊と創造の両方で雇用市場に影響を与える能力をもつ。この点で、創造性の発揮のような人間が発展させている特定のスキルは、将来ますます重要になるであろう。人間が AI システムと一緒に仕事をする仕事の数は増加し、それが新しい職場環境をもたらす。

また、AI は商品とサービスを、ときには隠れた自動化されたやり方でユーザーに推薦することができるので、AI が消費者の意見に積極的に作用し影響を及ぼす可能性があることは明らかである。

6.1.1 意思決定の変化

AI は、特に定型的なプロセスにおいて意思決定への関与を増すことが期待されている。正しく訓練されていても、技術は誤りを犯す可能性がある。問題は、人間と AI システムがどのように協調し、十分に熟慮して決定を行えるかである。しかし、AI と人間の推論のそれぞれの長所を、どのように適切なやり方で組み合わせられるのかはまだ明確でない。開発者にとって、完璧なアルゴリズムを作るという課題は非現実的なものである。アルゴリズムが複雑になるほど、それが社会又は産業に与える影響は大きくなり、意思決定において適切な質を確保するために人間の判断がより必要とされる[13]。

このことは、意思決定の責任に関して重要な意味をもつ。複雑なシステムでは、人間と AI システムのどちらに責任があるのかを区別することが難しい。AI は、依然として統計的な相関に頼っているが、意思決定に責任をもっている印象を与えることがある。このため、意思決定者は責任を負うことを拒み、AI システムに責任を押し付ける可能性がある[13]。

6.1.2 先進的なサプライチェーンの運用

AI を使用すると、発注をするためにカタログと価格見積もりをただ次々に見ることに労力を費やすという作業は、過去のものとなるであろう。代わりに、発注者は自らが希望する品物の写真を撮るだけでよく、その後、その写真は供給業者の有効在庫と自動的に照合される。発注者が全く新しい品物を注文しようとする場合には、その要求事項を簡単に説明しさえすれば、それだけで、AI が価格提示を要求し、供給業者の見積もりをアセスメントし、価格、品質及び納期に応じて最も適切な品物を選択するのに十分で

ある。特定のコンポーネントが必要だが、その部品番号又は仕様のようなその品物を再注文するのに必要な重要な情報が欠けている場合には、品物を特定し、適切な交換を提案するためにも AI を使用できる。

これらの全てを AI システムが決定し始めると、一部の市場は近い将来、深刻な混乱に陥るであろう。消費者はもはや実店舗を訪れず、ワンクリックで全ての製品を手に入れる。サプライチェーンが変化すると、中間業者は関与を失う。消費者の選択がますます AI アルゴリズムに影響されるようになるにつれ、推奨の透明性がより大きな課題になる。例えば、スマート冷蔵庫が食品の供給業者と冷蔵庫の所有者の食事に関して自律的に決定を行うと、それは食料品店など市場全体を混乱させる可能性がある。また、消費者と市場参加者は、AI システムが推奨する供給業者による不正行為から保護される必要がある。

6.2 データ関連の課題

AI は、機械学習アルゴリズムの訓練のために大量のデータを必要とする。しかし、データの共有と配布は、適切な規則及び規制がないために現在様々な産業部門で制約されている。このため、様々な産業の実務者が自らの商業的利益に従ってデータを隔離し、データ共有に関して境界を設定している。これらの産業での、かなりの量のデータの蓄積にもかかわらず、データアイランドの問題が AI の可能性を完全に実現する上で障害になっている。

加えて、全体的なデータの利用可能性の問題が、機械学習アルゴリズムのために信頼性の高いデータを十分に収集することを困難にする。これらのアルゴリズムは人間の脳とは異なる働きをし、一般に、統計的にまれである内容を考慮することなく、データ内容を基本的な真実として扱う。これは、AI 技術を適用するときに認知の偏りと誤解をもたらし得る。特定の問題を解決しようとする際、集めたデータの総量の中での信頼、信用のできるデータの割合が、機械学習の要求事項の下限に達しないことがあり得る。

6.2.1 訓練データの選択

訓練データ自体に偏りがあるせいで AI システムが性別又は人種の偏りを示しているケースが、メディアで注目を集めることが多くなってきている[96]。AI モデルを開発するときは、その開発で使用するデータの関連性、量及び質が重要である。

第一に、利用可能なデータは、当該問題の解決に関連したものでなければならないが、これは関わりのあるデータセットのサイズを考えると、判断することが必ずしも容易ではない。したがって、データ科学者は、データがどのように、何の目的で使用されるのかについてある種の理解を必要とする。データの関連性をアセスメントすることは、データ科学者だけでなく当該分野の専門家を含む学際的な作業となっている。第二に、モデルが正確に機能するために、モデルがデータセットの一般化を学べるのに十分なデータが存在しなければならない。そして第三に、訓練データセットの質は、モデルが採用された後にそれが扱うデータを代表するようなものでなければならない。

これらの条件が満たされているかどうかは、多くの場合、異なる様々なモデルで初期の訓練と試験を数回行った後に、初めて確認できる。このプロセスは非常に反復的であって、データ科学者はビジネスに対する自らの理解に依拠すると共に、モデルのパフォーマンスを検証するためにデータを試験しつつ、訓練データとモデルを数回調整する必要がある。試験データが元のデータセットから適切に分取されていない場合には、結果として試験データ自体が問題をもたらす可能性がある。訓練前に、データセットは、訓練のためにモデルに対して示すデータと、訓練中にモデルに示すことはせずモデルの質の試験のために使用するデータとに分割される。訓練データが確実にモデルの質を正確に試験できるようにするため、試験データが訓練データと同様に現実世界の状態を模擬するように分割することが望ましい(例えば、予測モデルの場合、試験の各データ点は、訓練データの最後のタイムスタンプの後のタイムスタンプをもつことが望ましい)。

モデルがほぼ常に正確に機能して期待を満たす、又は期待を上回る場合であっても、さらに時間を費やして、モデルが犯した誤りがアプリケーション全体のユーザビリティを損ねるほど深刻でないことを検証することが重要である。例えば、モデルが事例の98%において画像を正しく分類するが、画像の0.2%を偏った有害なやり方で間違っただけで分類する場合には[97]、誤り率は全体として非常に低いにもかかわらず、そのモデルを採用することはできない。

こうした事例は、モデルが適用のために学習した訓練データ自体に、人間的な偏りが存在することの結果であることが多い[98][99]。偏りが生じがちな属性(人種、性別、性的指向又は宗教など)を訓練データから削除することさえ、他の変数がそれらを代理するように作用し得るため、モデルからそうした偏りを除去するのに十分ではないかもしれない。偏りを抑制するための技術的方法は存在するが、どれも完璧ではなく、より洗練されたアプローチを開発するにはさらなる学際的な研究が必要である[100]。

データ収集量が増加するにつれ、それらがより代表的で偏りが無いものにするのも、この問題の解決に向けた重要な最初のステップになる。それまでは、訓練データとモデルの偏りをチェックして正す方法を開発するためにさらなる作業が必要である。

6.2.2 標準化データ

既に述べたように、AIの成功は使用するデータの量、種類及び質に大きく依存する。現在のデジタル変換の過程では、既に様々なチャンネル(例えば、リンクされたオープンデータ、多数のセンサ、既存のデータベース)を通じて、大量のデータがアクセス又は生成される可能性がある。多様性は恵みであるが、課題でもある。適切な理解のためにデータを前処理して記述すれば、分析結果を大きく改善することができる。将来、この時間を要するステップは、データ型、フォーム及び情報モデルを標準化することで能率化されるであろう。

ここで生じる問いは、異種の情報とデータセットを、特に幾つかのAIアプリケーションに渡すそれらを、関連するデータセットの意味をまず発見することなしにどのように適切に理解し、解釈できるかである。これを達成するには、製造業者に依存しないあいまいさのないデータの情報モデルが必要である。セマンティック(意味論的)技術は、機械と人間が理解できる情報の統一的な表現を確実にし、データを明確で包括的な形で利用できるようにすることが十分に証明されている。これに基づいて、適切なセマンティックツールは暗黙の知識を引き出すことを容易にし、それ自身が、ある形態の効率的なデータ前処理を表している。

したがって、セマンティックな相互運用性は、システムが関連データを交換する、又は以後の処理でそれを利用できるようにすることを要求するだけでなく、交換されるデータの送り手と受け手による解釈が同一であることを要求する。例えば、同一のデータ点異なる用語で記述されている、又は異なるデータ点同一の用語で記述されている場合には、セマンティックコンフリクトが発生することがある。

しかし、異種データの理解は、意味的にだけでなく構文的にも保証され、標準化されなければならない。データセットの構文的な相互運用性は、データが交換される構造とフォーマットが十分に定義されていることを意味する。例えば、二つのシステムが、関連データを提供又は処理するために異なるフォーマット又は構造を使用すると、構文的コンフリクトが生じる。異なるレベル及び通信チャンネル上での、標準化された交換フォーマットと通信プロトコルがこうした障害を解決する。

6.3 アルゴリズム関連の課題

AI で使用されるアルゴリズムに関連した課題も存在している。これらのアルゴリズムの採用における最も注目すべき幾つかの問題は、ロバストネス、新しいタスクに適応する能力、及びそれらの解釈可能性の欠如である。

AI アルゴリズムの安全性とアルゴリズムから生じるユーザーのリスクは、複雑なシステムの場合にますます重要な課題になっている。アルゴリズムが正しい動きをすること、すなわち、あらゆるデータに対してアルゴリズム又はプログラムがその仕様に記述されている問題を正しく解決することを確実にしなければならない。これは機械学習アルゴリズム、特にニューラルネットワークにとって非常に大きな課題であり続けている。例えば、多くの場合、ニューラルネットワークがどの問題を解決する必要があるかに関して明確に定義された仕様は存在しない。加えて、アルゴリズムの複雑さが意思決定プロセスを理解することが困難に、又は不可能にさえしている。これらの課題の幾つかを以下、より詳しく概説する。

6.3.1 ロバストネス

機械学習に関連するロバストネスという用語は、入力が訓練データと異なってもアルゴリズムが正しい決定をすることを意味する。したがって、ロバストなアルゴリズムは、敵対的な入力に対して安定しており、訓練データセットとアプリケーションデータセットの間にパフォーマンスの点で大きな偏差がない[101][102][103]。

ロバストネスは過去にも既に課題とされていたが、今後のトレンドがその重要性を高めるであろう。特に強化学習の成功が、ロバストなアルゴリズムの研究の進歩に貢献している。4.1.3 節に記述したように、このカテゴリの機械学習アルゴリズムは、自らの環境及び他のエージェントと相互作用するエージェントを使用する。これは相互作用と変化する環境の非常に複雑なシステムをもたらす、それが結果とエージェントの行動の予測を困難にする[101]。

意思決定のためにアルゴリズムを使用するトレンドは、おそらくロバストネスに関して最大の効果をもっている。これらの決定が大きな影響をもつほど、現実環境で正しく応答するアルゴリズムの能力がより重要になる。

二つのシナリオが意思決定に関係している。すなわち、システムがそのオペレータにある決定を推奨し、オペレータがその推奨をレビューして検証するか、又はシステムが自動的に自らの決定を実施するかである。後者は、新しい入力データの場合に、システムが自らが誤りを犯したことに必ずしも気づかないという問題をもたらす。そのため、決定は検証に付されることがなく、インフラストラクチャ、又は人間にさえダメージを与える可能性がある。これにより発生し得るのは、例えば、癌患者のスキャン結果を誤って分類してしまい、その患者が適切な治療を受けられなくなる、あるいは機械学習システムが電力グリッドに過負荷を与えて事故を生じさせるといったことである。

こうした状況での AI アルゴリズムの失敗の幾つかの理由は、ミスマッチなデータセット、異常値、及びシステム自体のプログラミングである。現実のデータと一致しないミスマッチなデータセットは、正しく機能しないアルゴリズムをもたらす。例えば、異常値などに起因して大きな違いが存在する場合には、パフォーマンスも低下する可能性がある[102]。アルゴリズムは、望ましい出力からあまり逸脱することなしに、データセットの変化に適応できる必要がある。

アルゴリズムのロバストネスの分野で、現在、次のような幾つかの研究方向が着手されている。検証(すなわち、正しく機能するシステムをどのように構築するか)、妥当性(すなわち、システムが正しい目的を満たし、望ましくない行動をとらないことの保証)、セキュリティ(すなわち、第三者がシステムを操作することをいかに防ぐか)、及び制御(すなわち、最終的には人間がシステムを制御でき、問題が生じたらシステムを修正できることが必要) [104]。AI システムを訓練するためのデータの前処理、ミスマッチと異常値の除去、変化と異常の検出、及び仮説検定と転移学習など、複数のアプローチによって AI アルゴリズムのロバストネスを向上させることができる[102]。

6.3.2 転移学習

機械学習の実施は現在、カスタマイズされた製品になっている。変数は、問題に正確に合うように選び、訓練データはアプリケーション分野から得なければならない。これにより、アルゴリズムは確実にそのアプリケーションに対して完全に機能するようになる。人間は新しい問題を解決するため、それらに対して過去の知識を転用することができるが、機械はこの能力を持っていない。諸変更が提供されたデータの分布に影響を与え、それがデータ又は類似のアプリケーション分野をもはや古いものにしてしまう場合には、訓練データを再び収集し、アルゴリズムを再び訓練する必要がある。転移学習は、コストと労力を減らすために、知識をあるアプリケーションから別のアプリケーションへ移すことに役立つ[105]。

図 6-1 に示すように、転移学習の目的は、異なった分布又は異なったタスクをもつ異なったアプリケーション分野からの訓練データを使用できるようにすることである。どの知識を移すことができるかを判断する方法など、幾つかの問題に対処する必要がある。訓練データが別のアプリケーション分野からのものである場合、移された情報は関連性が有ること

も無いこともある。このアプローチに関しては、知識をいつ、どのように移せるかなど多くの問題がまだあるが、改善が今なされている。転移学習でブレイクスルーが起きれば、機械学習の適用ははるかに容易になり、開発のコストと時間を削減できる。

ロボットや自律走行車のように、システム全体が人間との相互作用に関して安全な必要がある場合には、アルゴリズムのロバストネスは不可欠な要因である。

6.3.3 解釈可能性

ほとんどの AI アルゴリズム、特にニューラルネットワークは“ブラックボックス”として記述される。これは、入力データ及びネットワークの結果は理解できるが、アルゴリズムがどのようにその結果に達しているのかは理解できないことを意味する。モデルを理解することがエンドユーザーに広く受容されるための最も重要な出発点の一つであるため、これは AI にとって重大な課題である。回帰又は決定木など一部のモデルはデータ科学者と AI の専門家にとって理解可能であるが、データフローの次元、及びその他のほとんどのアルゴリズムの複雑さは通常、高すぎて適切に理解することができない。

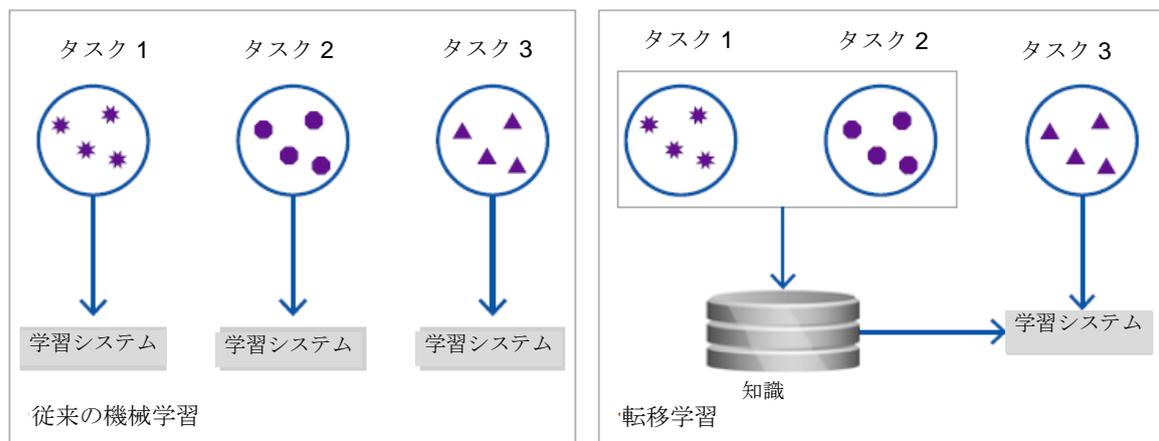


図 6-1 | 従来機械学習と転移学習の比較

このことは、特定の予測をなぜ行ったのかに関して、こうしたアルゴリズムが明確な説明を与えないことを意味する。アルゴリズムは単に確率を示すだけであり、それはしばしば解釈することが難しい。これは訓練したアルゴリズムが期待通り機能することの検証を難しく、及びしばしば不可能にする。ときには何百万ものモデルパラメータが関わっていて、入力する特徴とパラメータの間に1対1の関係がないことがある。したがって、多くのパラメータを使用し、複数のモデルを組み合わせることは、しばしば予測に影響を及ぼす。これらの中には、高い精度を得るのに大量のデータを必要とするものもある。

しかし、問題を生じさせるのはアルゴリズム自体だけではない。生データを、数学モデルで処理することができるデータに変換すると、単純なアルゴリズムさえ、人間が解釈できないものになり得る。幾つかの方法が、NLP 又は画像認識のようなアプリケーション分野でニューラルネットワークを解釈可能にするために提案されている。他のアプローチは、複雑なアルゴリズムを単純で理解可能なモデルによって局所的に近似することで、より解釈可能にすることを試みている[106]。

AI アルゴリズムが、医学又は金融のような極めて注意を要するデータを伴う分野を含め、ますます多くの分野で採用されるにつれ、解釈可能性は確実に重要性を増し、これは今後、AI の大きな概念的及び技術的課題の一つであり続けるだろう[107]。AI アルゴリズムの十分な検証と妥当性確認は依然として問題が非常に多く、それらの影響は、アルゴリズムの目的関数の説明を通じて示すことができる。

6.3.4 目的関数

ここで主に注目するのはAI システムの目的関数であり、これが誤っている又は不正確であると、負の副作用又は報酬ハッキングにつながる可能性がある。負の副作用は、例えば、システムが商品又は人間に対して害をなすことを含み得る。これは、そうした誤った又は不正確な形での動作が、より急激に目的を達成することを許してしまうためである。一方、報酬ハッキングは、システムのタスクを不適切な形

で完了させる。これはシステムが、その報酬関数を満たす上で、予測外の代替方法を発見してしまったためである[106]。システムの目的関数はまた、システムの停止又は変更を妨害しないことが望ましく、これは自らの目的を達成するシステムの能力に影響が及ぶ場合でさえそうである[106]。

目的関数が正しく記述されているときでも、システムは、スケーラビリティの問題がその管理の範囲内にある場合に、例えば、システムのパフォーマンスについて頻繁にフィードバックを行うと費用がかかり過ぎる場合に、正しく機能できる必要があるであろう[108]。

前述のように、不適切な訓練データは非常に望ましくない結果をもたらす可能性がある。したがって、対処する必要がある問題は、システムが訓練段階で遭遇しなかった新規のデータに遭遇したとき、システムがどのように作動することが望ましいのか、例えば、誤った攻撃的な動作をするよりは、むしろ穏やかな失敗をするといったことである[109]。不適切な訓練データ以外に、強化学習エージェントのための試験と訓練環境も、それらを経験することで発生し得る悪影響を封じ込めるために、安全で隔離されている必要がある[101]。

6.4 インフラストラクチャ関連の課題

AI アプリケーションを良好なパフォーマンスで(特にリアルタイムという制約のもとで)実行するために、計算速度とインフラストラクチャの効率を着実に高める必要がある。AI の作業負荷を高速化するようにカスタマイズされたハードウェアが必要なだけでなく、基礎となるリソースを最適に活用してプラットフォーム上でAI タスクを実行することを可能にするソフトウェアスタック、ライブラリ又はツールチェーンも必要である。

6.4.1 ハードウェアのボトルネック

AI とディープラーニングは特に大量のデータの並列処理を必要とし、従来のコンピューティングアーキテクチャがこれを担うことはほぼ不可能である。現在使用されている GPU と FPGA は、最先端の AI アルゴリズムの実行を制約するような幾つかの技術的

制限をもっている。例えば、ディープラーニングに最初に導入された GPU は、次のような三つの主な制限をもっている。これは、並列計算の長所を十分に活用することができない。プログラミングが可能でなくハードウェア構造が固定されている。ディープラーニングのアルゴリズムの有効性にまだ改善の余地がある。新しいコンピューティング時代にあつては、コアチップが AI のインフラストラクチャとエコシステムを決定する。したがって、プロセッサの能力は AI 開発を進める上で大きなボトルネックと考えられている。

そうしたことから、異種コンピューティングプラットフォーム(これは多様な AI 作業負荷を扱うために様々なアクセラレータを統合する)の設計とアーキテクチャは、AI の研究と商業的な実施に欠かせない事項である。さらに、クラウドインフラストラクチャ内で提供されるハードウェアリソースは、それらのスケラビリティ、信頼性及び自動化されたリソース管理ゆえに、最近のトレンドになっている。また、例えばコンテナのような、クラウドネイティブのアプリケーションプログラミングインターフェイス(API)が、一貫性のあるインタフェース、幅広いサポート、及び AI アプリケーションの採用の容易さのために使用されている。

AI 技術は様々なシステム又はサブシステム(クラウド、エッジ、又はエンドデバイス)上で実行され得るので、プラットフォームの設計は、システムの個々のニーズとリソースの制約に合わせて調整することが望ましい。例えば、クラウドサーバは、モバイル機器よりも高度なアルゴリズムを実行し、より大量のデータを処理する可能性がある(例えば、モデル訓練のために)。したがって、ハードウェア設計では、様々なシステム又はサブシステムにおける AI 機能の調整を考慮に入れる必要がある。

6.4.2 プラットフォームとフレームワークの欠如

AI の開発のための、再利用可能で標準化されている技術的フレームワーク、プラットフォーム、ツール及びサービスはまだ成熟に至っていない。少数のオープンソースの AI 学習システムとディープラーニングライブラリは、よく知られた技術界の巨大企業から得られるようになっているが、アーキテクチャ、フレームワーク、アプリケーションモデル、アセスメントと視覚化のツール、及びクラウドサービスの、完全にモジュール化され、標準化された AI エコシ

テムが、適切な成熟レベルに達するにはまだ時間を要するかもしれない。

6.5 信頼性関連の課題

AI は、協力して共に仕事を進める必要がある様々なステークホルダーが関わるトピックであることが、広く認識されている。例えば、製造に関する予測保守の分野では障害はまれにしか発生しない。AI アルゴリズムを適切に導入するには、製造業者とユーザーの両方がデータを共有し、専門知識を提供し、効率的な実施に向かって協力する必要がある。この協力を容易にするために、信頼の確保など、幾つかの問題に対処する必要がある。

6.5.1 信頼

機械学習アルゴリズムは、与えられるデータに依存している。したがって、意思決定の自動化には完全で正確なデータが不可欠である。データ品質の悪さのような潜在的問題、又は意図した操作さえが、価値のない結果を発生させ、さらにはアルゴリズムのユーザーに悪影響を及ぼす可能性もある。

関係者の間の信頼は不可欠なものである。データソースの信頼性の問題に対処する解決法が、認証技術から与えられる可能性がある。一元化された信頼できる発行者からの電子証明書と、データのシーリング(封止)との組み合わせが、当事者間で信頼を確立するための選択肢である。しかし、この解決法はパートナー間の信頼を確立することだけを狙いとし、データの質の問題には対処していない。この目的に関しては、信頼できるデータプール(集積)を集めるか又は誤ったデータベースを回避するための、評価又はアセスメントのアルゴリズムを使用することができる。メタアルゴリズムは、時間が経過しても AI システムの信頼性と透明性が維持されることを助け、使用した情報ソースの出所と配布に関する情報を与える[13]。

6.5.2 プライバシー

AIの開発は、データ訓練アルゴリズムの使用に依存している。このプロセスでは、大量のデータを収集、分析及び使用する必要がある。データの価値はますます顕著になってきている。開発者、プラットフォームのプロバイダ、オペレーティングシステムと端末の製造業者、及びバリューチェーンのその他のサードパーティは、これらのデータにアクセスし、ユーザーが提供したデータのある程度までアップロード、共有、変更、交換及び活用することが可能である。

さらに、AIシステムは一般に高いコンピューティング能力を必要とするので、多くの企業及び政府はクラウド上にデータを保存し始めている。しかし、クラウドのプライバシー保護にも、隠れた脅威が存在している。合法的に、現在と将来の法律に従っていかにデータを収集し、使用するかは、全てのAI当事者にとって重大な問題である。

6.5.3 セキュリティ

技術の悪用、欠陥、及び将来のスーパーAIの開発は、いずれも人間社会にセキュリティ上の脅威をもたらす。AIが人間に与える影響は、人々がそれをどのように使用し、管理するかに大きく依存する。AIを犯罪者が手にすれば、大きなセキュリティの問題が確実に生じ得る。例えば、ハッカーは、AIシステムのユーザーの挙動を自ら学習して模倣できるようなソフトウェアによりサイバー攻撃を行い、可能な限りシステム内に留まるために、絶えず方法を変化させるかもしれない。技術的な欠陥の中には、異常な作業を発生させてAIシステムをリスクにさらすものもある。例えば、ディープラーニングに使用されるブラックボックスモデルは、モデルを解釈不可能なものにする。したがって、不適切な設計は異常動作につながる可能性がある。さらに、セキュリティ対策が十分に効果的でない場合には、無人運転車、ロボット及びその他のAI機器が人体に害を与え、これは法律的な観点からの課題になる可能性がある。

6.6 規制関連の課題

適切な規制が、多くのAI分野でまだ欠如している。産業革新、生産性及び競争力を促進、支援する一方で、高いレベルのセキュリティと健康、消費者保護、社会保障と権利及び自由の保護を同時に確保するような、AI開発に対するバランスの取れた規制のアプローチを発見することが、世界の多くの政府にとって重要な優先事項である。

無人運転車及びドローンなどの分野では初期の立法的措置が少しは講じられているが、AIに特化した規制機関は世界のどこにも存在せず、AIに関する法的な調査も欠けている。例えば、欧州では、ロボティクスとAIの領域は、国及び欧州レベルで様々な規制当局及び機関がカバーしている。これらの領域において技術的、倫理的、規制的な専門知識を提供し、開発を監視している欧州の中央機関は存在しない。このような調整の欠如は、これらの技術開発から生じる新しい機会及び課題に対してタイムリーに、十分な情報に基づいて対処することを妨げている。

欧州議会の法務委員会のAIに関する報告書で特定されている、六つの主な領域横断的規制テーマは、幅広い政策分野に関係している。委員会の立場から、優先的に行動が必要とされている分野は、自動車部門、高齢者の介護、ヘルスケア及びドローンを含んでいる。

6.6.1 責任

新しいAIベースの製品と共に生じてきている予見可能性、解釈可能性、及び因果関係の問題は、製品欠陥などの責任問題に対処することをますます困難にし、これは大きな責任ギャップを生み出し得る。予想されるこれらの責任の問題に直面することで、新しい規則と規制の必要性、例えば不法行為及び契約の法律におけるそれが、多くの産業でますます重要になってくる。責任に関する法的安定性は、イノベーター、投資家及び消費者にとって最も重要な事柄であり、彼らに必要な法的フレームワークを与える。

しかし、デジタル技術の複雑さは、誰に責任がある

のか、また障害が発生した場合にどの程度責任があるのかの判断を特に難しくする。例えば、既存の法的カテゴリは、ロボットの法的性格を適切に定義し、そこから損害に対する責任を含む権利と義務の帰属先を決めるには十分でない。現在の法的フレームワークの下では、ロボットは、第三者に損害を生じさせる行為又は不行為に関して自身が責任を負うことはできない。ロボットが自律的な決定を行えるシナリオにあっては、伝統的な規則では補償を行う責任のある当事者を特定し、生じた損害の補償をその当事者に要求することができないため、そうした規則はロボットに責任を課すには不十分である。

欧州連合では、欧州委員会(EC)が、欠陥製品があった場合に消費者と生産者に関して法的明確性を確保するために、技術的發展に照らして製造物責任指令の解釈の手引を 2019 年半ばまでに発行すると予想される。

6.6.2 プライバシー

欧州連合の GDPR(一般データ保護規制)などの規制は、部分的にはこれらの問題に対処することを意図しているが、データ保護当局がこれをどのように実施するかはまだ不明である[24]。データのプライバシーと、AI 産業界の繁栄を実現することとの間で、微妙なバランスをとる必要がある。実際、近いうちに AI 自体が、洗練された匿名化と暗号化の方法を実現することによって、個人データの安全性を確保することを助けるであろう。フェデレーションラーニングが、AI システムの訓練のために個人データが消費者のデバイスから流出する必要がないことを保証する。これは、システムが全てのデバイスにおいて直接並行して訓練されるためである[110]。さらに AI は、人間によるデータへのアクセスを必要とせずにタスクを実行することによって、扱いに注意を要する情報(例えば健康記録)の露出を制限し、それによってプライバシーを高めることができる。

GDPR は、欧州連合でのデータ保護とプライバシーに関する一連の重要な規制変更であり、AI システムによる自動化された意思決定にも対処している。具体的には、GDPR は、“当人に関係する法的効力発生

させる又は当人に重大な影響を及ぼすプロファイリングを含め、(当人の明示的な同意なく、あるいは欧州連合又は加盟国の法律第 22 条(2) GDPR [111]による承認なく)自動処理だけに基づいた決定の対象にはならない権利”を人に与える。GDPR はまた、“管理者サイドの人間の介入が得られ、当人の見解を表明でき、決定に異議を唱えることができる権利”、第 22 条 (3) GDPR [111]も人々に与える。

さらに、GDPR の第 13 条～第 15 条は、第 22 条に該当する自動意思決定が行われる場合には、人々がそのことを告げられ、アルゴリズムの基礎になっている決定プロセスについて意味のある情報を与えられ、自動化プロセスの結果及びその重要性について知らされることを要求している [112]。

データ保護当局が GDPR を実際にどのように実施するかは不明確なままであるが、AI の意思決定についての透明性の要求事項は、企業と規制当局、両方が対処する必要がある主要課題になる可能性が高い。ここでの目的は、データのプライバシーと透明性を維持することと、データ主導のビジネスモデルの繁栄を実現することの間でバランスをとることである。健全な AI エコシステムの実現は、単に経済的な観点からだけでなく、企業が透明性を確保する能力を向上させ得るさらなる技術研究の実現のためにも必要とされる。

6.6.3 倫理

これらの問題からの最も深刻な影響は、より先進的、未来的な AI システムにあってだけ見られるものだが、可能な限り早くそれらに対処する事前対応的なアプローチは、慎重なアプローチであるだけでなく、将来、費用のかかる(不可能ではないにしても)後付け対策を回避することにつながるかもしれない。

より急を要する関心事は、意思決定プロセスにおいて倫理的な選択をする AI システム(例えば、自動運転車)の必要性である(例えば、歩行者を負傷させるのか、あるいは歩行者を避けることで運転者又は乗客を負傷させる可能性を生むのか)[113]。この例は、AI の安全性が単なる技術的な問題ではなく、そうした

技術のユーザー、技術を中立的に傍観する人々、及び技術を開発する企業を保護するための学際的アプローチが必要とされるような、政策的及び倫理的問題でもあることを示している。これは、後者の場合、重要な法律的課題に直面することがあるためである。研究機関と企業がこうした問題に取り組み始めてはいるが、国際レベルにおける関係者全てのより緊密な協力が必要とされている。

AI は、幾つかの意思決定プロセスにおいて徐々に人間に置き換わっている。インテリジェントロボットは、決定を行う際に人間社会の倫理的な制約と規則に適合する必要もある。例えば、無人運転車が、ブレーキが間に合わない状態で正面の歩道には歩行者が 3 人いると仮定する。システムは、これら 3 人の歩行者に衝突することを選ぶべきか、あるいは代わりに方向をそらし、道路の反対側にいる 1 人の歩行者に向かって進むべきか？ 人間の日常生活における AI の適用は、対処する必要がある基本的な倫理的課題の中心に位置している。AI システムの設計が倫理的及び社会的制約と一致しない場合、そうしたシステムは人間の論理と異なる論理に従って動作し、劇的な結果をもたらすかもしれない。

さらに、機械に意思決定の権利を与えた後には、人々は新しい倫理的問題に直面するであろう。インテリジェントシステムが特定の分野で知識を獲得するにつれ、その意思決定能力は人間の能力を上回り始める。これは、人々がますます多くの分野で機械主導の意思決定に頼るようになり得ることを意味する。将来の全ての AI 開発において、この種の倫理的課題に早急に特別な注意を払う必要がある。

第7節

人工知能における標準化ギャップ

標準化は、AI 開発において支援的及び主導的、両方の役割を果たす。産業革新の促進だけでなく、AI 製品とサービスの質を向上させ、ユーザーの安全を確保し、公正で開かれた産業のエコシステムを構築することも不可欠である。

ここまで、現在の AI の展望とその主要課題をレビューしたのを受けて、標準化のための基本的な要求事項を幾つか導出することができる。この節では、まず AI に関連する既存の標準化の取り組みの概要を説明し、その後、標準化と産業のギャップの幾つかに焦点を合わせ、そこから次節の最終的な推奨事項を

導く。

7.1 人工知能における標準化の取り組み

AI の分野の標準化はまだほんの初期段階にある。AI 又はその支援技術は、ある側面についてはかなり前から既存の標準化グループの範囲の一部になっていたが、より広範囲で全体的な観点から AI 分野に対処するために、新しいグループが現在作られているところである。次節では、図 7-1 に示す組織を含め、現在の標準化状況の概要を説明する。

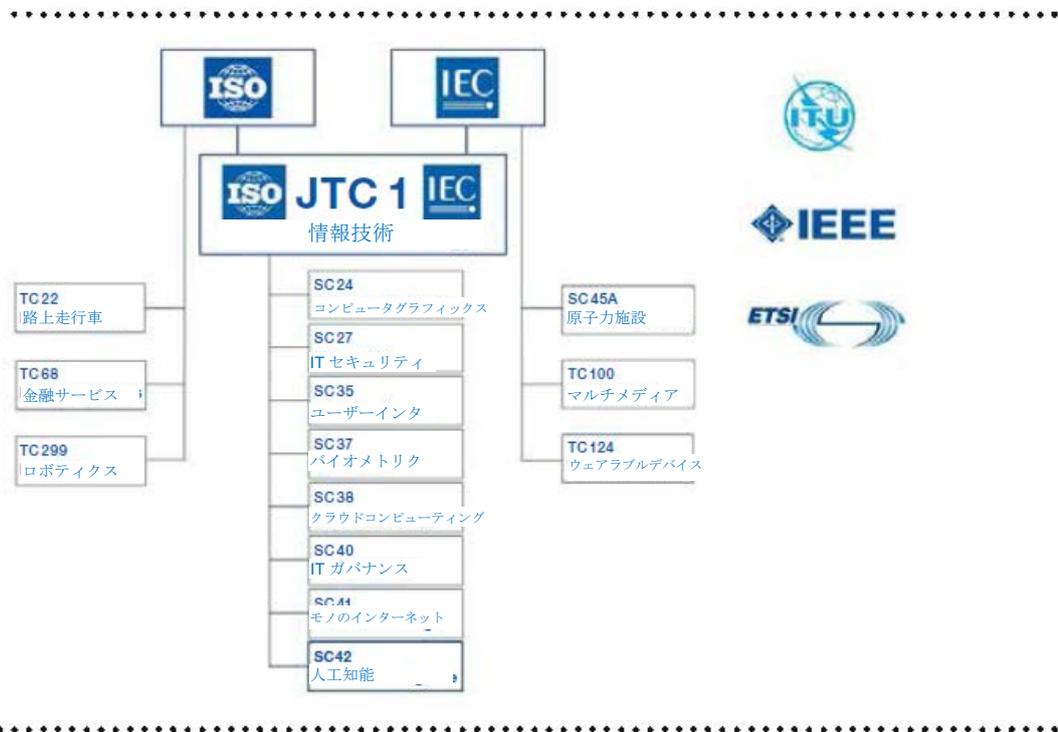


図 7-1 | AI に関する標準化の状況

7.1.1 ISO/IEC JTC 1

ISO と IEC の JTC 1 は、IEC と ISO の間で設置され IT の問題に関する合同専門委員会であり、AI 用語の分野で長期に渡って作業をしている。JTC 1 の以前の用語ワーキンググループが、AI 用語に関する次の一連の国際規格を発行している：

- ISO/IEC 2382-28:1995, 情報技術－用語集－第 28 部：人工知能－基本概念及びエキスパートシステム
- ISO/IEC 2382-29:1999, 情報技術－用語集－第 29 部：人工知能－音声認識及び合成
- ISO/IEC 2382-31:1997, 情報技術－用語集－第 31 部：人工知能－機械学習
- ISO/IEC 2382-34:1999, 情報技術－用語集－第 34 部：人工知能－ニューラルネットワーク

これらの古い部は現在、IT 用語に関する共通 JTC 1 規格 ISO/IEC 2382:2015 [11]に統合されている。

ISO/IEC JTC 1/SC 42

この分科委員会は、AI の特定の標準化要求事項に対処するために 2017 年 11 月に設立された。JTC 1/SC 42 の作業範囲は、AI に関する JTC 1 の標準化プログラムの中心的存在及び提案者として役割を果たし、AI 関連のアプリケーションを開発している JTC 1、IEC 及び ISO の委員会に手引を提供することである。この分科委員会の作業の一部となっているトピックは次の通りである：

- 基本規格
- AI の計算アプローチ及び特徴
- 信頼性
- 使用事例とアプリケーション
- ビッグデータ
- 社会的懸念

2018 年 4 月に開催された最初の本会議の後、JTC 1/SC 42 は基本規格に関する WG 1 を設置した。これは、承認された最初の二つのプロジェクトである、人工知能の概念及び用語(ISO/IEC 22989)及び機械学習を使用する人工知能システムのフレームワーク(ISO/IEC 23053)を含んでいる。

JTC 1/SC 42 の本会議で、次の三つの研究グループが設置された：

- SG 1：AI システムで使用される様々な技術(例えば、機械学習アルゴリズム、推論)を、それらの特性と特徴を含めて研究するための、人工知能システムの計算アプローチ及び特徴。それらの基礎にある計算アプローチ、アーキテクチャ及び特徴を理解し、明確化するための既存の特殊な AI システム(例えば、コンピュータビジョン、NLP)。AI システムのアプリケーションに関する産業界の慣行、プロセス及び方法。
- SG 2：信頼性(透明性、検証可能性、説明可能性、制御可能性などを通じて AI システムへの信頼を確立するアプローチの調査を目的としている)。工学的な落とし穴と、AI システムに付随する一般的な脅威とリスクの評価、及びそれらを軽減する技術と方法。AI システムのロバストネス、復元力、信頼性、正確性、安全性、セキュリティ、プライバシーなどを達成するためのアプローチ。AI システムの偏りの発生源の種類(そうした偏りの最小化を目的としており、AI システム及び AI が支援する意思決定における統計的偏りを含んでいるがそれに限定されない)。
- SG 3：様々な AI アプリケーション分野(例えば、ソーシャルネットワーク、組み込みシステム)、及びそれらが使用される様々な状況(例えば、ヘルスケア、スマートホーム、自律走行車)を明確化することを目的とした、使用事例及びアプリケーション。代表的な使用事例の収集。ISO/IEC 22989 及び ISO/IEC 23053 プロジェクトで定義されている用語と概念を使用して、アプリケーションと使用事例を説明し、必要に応じて用語を拡張。

その他の JTC 1 分科委員会

AI は他の多くの IT 分野とアプリケーションに影響を与える横断的な技術であり、他の JTC 1 分科委員会は次のように、推進要因及び支援技術として AI に関わりをもつ標準化の作業を行っている：

- JTC 1/SC 24 : コンピュータグラフィックス、画像処理及び環境データ表現
- JTC 1/SC 27 : IT セキュリティ技術
- JTC 1/SC 35 : ユーザーインタフェース
- JTC 1/SC 37 : バイオメトリクス
- JTC 1/SC 38 : クラウドコンピューティングと分散型プラットフォーム
- JTC 1/SC 40 : IT サービス管理と IT ガバナンス
- JTC 1/SC 41 : モノのインターネットと関連技術

7.1.2 IEC

幾つかの IEC 委員会は、AI を、彼らの作業プログラムに寄与する可能性がある一つの要素として見てきた。その例は下記を含んでいる：

- “SC 45A : 原子力施設の計測制御及び電力システム”は、原子力の計器と制御の要求事項をサポート及び調整する先進的なコンピュータと情報システムを進歩させるために、新しい IT 及び電子技術を適用することを目的として AI に関する研究を行った。
- “TC 100 : オーディオ・ビデオ・マルチメディアシステム及び機器”は、ウェアラブルデバイスに関する規格を開発し、AI と仮想現実の要素を含む“ウェアラブルデバイスの使用シナリオ”と題した議論トピックに着手した。
- “TC 124 : ウェアラブルエレクトロニックデバイス及びテクノロジー”は、ウェアラブル技術の電気工学、材料及び個人の安全に関する技術規格の策定を担当するために最近作られた。こうしたデ

バイスと技術は、AI アプリケーションの状況で広く使用されることが予想される。

7.1.3 ISO

幾つかの ISO 委員会が、次のように AI アプリケーションに関連する規格の作成に関係している：

- “TC 22 : 自動車”は、路上走行車の基本規格を策定し、インテリジェンスカーとコネクテッドカーに関係する標準化の課題も研究している。
- “TC 68 : 金融サービス”は、金融と銀行部門の標準化に取り組んでいる。新しいトレンドは、この委員会の約 58 の規格が扱っている。
- “TC 299 : ロボティクス”は、様々な用途のロボティクスの標準化の分野を扱っている。

7.1.4 ITU

AI の領域で、ITU-T は 5G (FG-ML5G) を含む将来のネットワークのための機械学習に関するフォーカスグループをもつ。このグループの目的には、次のものが含まれる[81]：

- アーキテクチャ、インタフェース、使用事例、プロトコル、アルゴリズム、データフォーマット、相互運用性、パフォーマンス、評価、セキュリティ及び個人情報保護を含む、将来のネットワークにおける機械学習の採用を支援する。
- 将来のネットワークにおける機械学習のために、既存の技術、プラットフォーム、ガイドライン及び規格を研究、レビュー及び調査する。
- 機械学習のフレームワークの安全で信頼できる使用を可能にする要因を明確化する。

- 将来のネットワークでどのように機械学習アルゴリズムを訓練、適合化、圧縮及び交換するか、及びどのように複数のアルゴリズムが相互作用するについてレビューし、研究する。
- 様々な固定及びモバイル通信スタックを考慮に入れて、将来のネットワークに適用される機械学習の考え得る要求事項を明確化し、こうした要求事項を満たせる新しい機械学習方法の開発を推進する。
- 機械学習を使用するためのネットワーク機能、インタフェース及び能力に関する考え得る要求事項を明確化する。
- 通信のための機械学習における標準化の課題を明確化する。
- これらのトピックに関する ITU-T 推奨の関連範囲を明確化するために、ギャップ分析と機械学習のロードマップの作成を行う。
- IEEE P7004 : 子供と学生のデータガバナンスに関する規格
- IEEE P7005 : 透明性のある雇用者のデータガバナンスに関する規格
- IEEE P7006 : 個人データ人工知能エージェントに関する規格
- IEEE P7007 : 倫理主導のロボティクス及び自動化システムに関するオントロジー(存在論的)規格
- IEEE P7008 : ロボットシステム、インテリジェントシステム及び自律システムの倫理主導の推進に関する規格
- IEEE P7009 : 自律システムと半自律システムのフェイルセーフ設計に関する規格
- IEEE P7010 : 倫理的な人工知能及び自律システムに関する福祉メトリックス基準
- IEEE P7011 : ニュースソースの信頼性を識別して評価するプロセスに関する規格
- IEEE P7012 : 機械可読個人プライバシーの条件に関する規格

7.1.5 IEEE

IEEEはこの領域で、主にAIに関連する技術規格の倫理的側面の研究に焦点を当てている。2016年3月、IEEE標準化協会は、AIがもたらす脅威に人々が対処することを支援し、データプライバシーからフェイルセーフ工学までに渡る倫理的な設計原則と規格を策定することを狙いとして、人工知能と自律システムにおける倫理的考察のための世界イニシアチブを開始した[82]。

この下で、IEEEはこれまで次の標準化プロジェクトを承認している。

- IEEE P7000 : システム設計における倫理的懸念に対処するためのモデルプロセス
- IEEE P7001 : 自律システムの透明性
- IEEE P7002 : データプライバシープロセス
- IEEE P7003 : アルゴリズムの偏りの考察

7.1.6 ETSI

ETSIは、経験的ネットワークインテリジェンス(ENI)に関する産業仕様グループ(ISG)をもっており、その目標は、閉ループ制御アプローチを含むコグニティブネットワーク管理システムに関する規格を開発することである。このアプローチは、“監視-分析-計画-実行”モデルに基づいており、学習能力によって高められて行く。

想定されているコグニティブネットワーク管理システムは、利用者のニーズ、環境条件及びビジネス目標のリアルタイムでの進展に応じて、利用可能なネットワークリソースとサービスの使用を管理できるようにする。システムが行う決定は、ネットワークリソースの複雑な状態についての詳細情報と、オペレータの好みを表している方針に依存している。

ISG ENI アプローチ独特の付加価値は、メトリックの導入によるオペレータエクスペリエンスの定量化、及び機械学習と推論の利用によるオペレータエクスペリエンスの最適化と調整である。

様々なAIメカニズムを使用して適応行動の変化を促すために、このグループは様々な種類の方針をレビューする予定である。ISG ENI は、リソース管理、サービス管理、オーケストレーション及び方針管理など、レガシーのネットワーク機能と発展中のネットワーク機能に関する既存の標準化されたソリューションを、該当する場合はレビューし、再利用する。

7.1.7 中国における標準化活動

- **国家情報技術標準化ネットワーク(SAC/TC 28)**

SAC/TC 28 は、用語集、ユーザーインターフェース、バイオメトリック特徴認識、及びその他の分野に関連したAIの標準化活動に主に取り組んでいる。

用語と用語集の領域では、GB/T 5271.28-2001: *情報技術—用語集—第28部:人工知能—基本概念とエキスパートシステム*など、これまでに四つの基本的な国家規格が発行されている。ユーザーインターフェース分科委員会が複数の国内規格を作成中であり、関連する標準化研究を実施するために、動作検知相互作用ワーキンググループと脳-コンピューターインターフェース・ワーキンググループを設置している。分科委員会は、国際提案“情報技術感情的(emotive)コンピューティング・ユーザーインターフェースフレームワーク”を提出し、これは既に承認されている。

バイオメトリック機能認識委員会は、指紋、顔及び虹彩の認識に関する規格を作成した。さらに、国家情報セキュリティ標準化専門委員会のビッグデータ規格に関するワーキンググループ、クラウドコンピューティング規格に関するワーキンググループ、及びセンサネットワーク国内規格に関するワーキンググループも、AIの関連技術とアプリケーションを支援するために基本規格の開発に取り組んでいる。

- **自動化システム及び統合に関する国内専門委員会(SAC/TC 159)**

SAC/TC 159 の下で、ロボット機器に関する SC 2 が、システムインターフェース、コンポーネント、コントローラなどを含む完全な形の産業用ロボットを担当している。SC 2 は、GB/T 17887-1999: *産業用ロボット—自動エンドエフェクタ交換システム—用語集及び特徴の提示*などの幾つかの規格を発表した。

- **オーディオ、ビデオ、マルチメディア及び機器に関する国内専門委員会(SAC/TC 242)**

SAC/TC 242 は、オーディオ、ビデオ及びスマートヘルスケア製品の関連規格の研究を行ってきた。現在の規格は、例えば、バーチャルリアリティオーディオの主観的評価方法(2017-0279T-SJ)を含んでいる。

- **情報セキュリティに関する国内専門委員会(SAC/TC 260)**

SAC/TC 260 は、AI 技術に焦点を合わせ、バイオメトリック機能の認識、スマートシティ及びインテリジェント製造などの分野におけるセキュリティ関連規格を策定した。

- **高度道路輸送システムに関する国内専門委員会(SAC/TC 268)**

SAC/TC 268 は高度道路輸送の領域で標準化活動を行った。また、GB/T 31024.2-2014: *協調高度道路輸送システム—短距離専用通信—第2部:媒体アクセス制御層及び物理層に関する仕様*などの規格も開発した。

7.1.8 米国における標準化活動

これを執筆している時点では、米国は現在のところAIに関連する方針又は規格をもっておらず、現在の政権にとってそれらを作ることが優先事項であるようにも見えない。AIを使用している米国に本社をもつ幾つかの民間企業は、幾つかの多国籍企業と共にAIに関するパートナーシップを結んでおり、これは

最良実施基準を開発して共有することを意図している。AIに関連する政策、倫理及び安全の問題に取り組む研究機関と非政府組織の数も増えている。政府には頼らない、AIに関するパートナーシップのような民間の協力イニシアチブが、予見し得る将来においては、米国を拠点とする企業にとって何らかの規格の源になる可能性が高そうに見える。

前の政権が政策と規格の開発により大きな関心を持っていたこと、また彼らがAIの課題と機会に関する広く引用されている二つの報告書を発行したことに注目すべきである。ただし、これらの文書は前政権の公式アーカイブの一部としてしかアクセス可能でなく、これを書いている時点では、それらが現政権によってどのように活用されるのか、あるいは活用されるのかどうかは、完全に明確にはなっていない。

7.1.9 欧州 AI アライアンス

AIは、ヘルスケア、農業又は行政など、欧州の重要な産業に影響を与えている。AIは欧州の産業、中小企業及び新興企業のビジネス機会も促進しており、欧州の生産性向上に貢献している。そのため、ECはAIの将来について議論するために欧州 AI アライアンスを締結している[80]。

ECは、信頼と認識の確保に関する合理性のある懸念に対処するための、関係する全てのステークホルダー(ビジネス、学者、政策立案者、消費者団体、貿易組織、及びその他の市民社会の代表者)を含む、AIに関する包括的な戦略の策定を約束した。結果として、ECはAI開発の全ての側面及びそれが社会と経済に与える影響に対処するために、複数のステークホルダーのフォーラムとして機能する欧州 AI アライアンスを締結することを計画している。

この欧州 AI アライアンスの設立に向けた第一歩は、AIに関するハイレベル専門家グループを作ることであり、これはアライアンスの作業の運営グループとして機能すると共に、中長期的なAIの課題と機会に関してECに助言するタスクをもつことになるであろう。この専門家グループは、他のイニシアチブと

の関わり及びそれに対する支援のメカニズムに関してECをサポートし、AI倫理ガイドラインを提案する。ハイレベル専門家の招集は2018年4月に完了し、グループは2018年後半に活動を開始することになっている。

7.1.10 コンソーシアム及びその他の組織

AIの規格と規則の問題は、一般の人々だけでなく企業、研究機関及び産業のコンソーシアムからもますます注目されるようになってきている。技術系の企業自身が、自らが開発しているAI製品とサービスの倫理的、経済的及び社会的影響、さらには必要とされ得る規格と規則について考察するようになってきている。

世界最大の技術系企業の幾つかは、これらの問いの一部に答え、自らの取り組みをモニターする助けとなる専門の倫理チームを作り、監査委員会を設置しさえしている。ほとんどの作業は現在、一般の人々の意識を高め、開発者と設計者に対する内部行動規則を策定することに集中している。このような取り組みは、万能のソリューションを生み出すことではなく、むしろどのような種類の規格と規則が最も良く機能するかに関する内部トライアルを通じて、十分な情報に基づいて企業としての意見を形成することを狙いとしている。

この種の内部作業は、他の様々なステークホルダーと連携するための貴重な基盤を築くことになる。ステークホルダーたちを引き合わせることを狙いとした、おそらく一般に知られている最も著名なイニシアチブは、“AIに関するパートナーシップ (Partnership on AI)”であり、これは50社を超える大手の技術系企業と、AIに焦点を合わせた研究機関による拡大しつつあるコンソーシアムである。その狙

いは、AI が確実に人々と社会に恩恵を与えようにすることである。このことは、AI システムの影響と設計に関する共同研究、最良実施法の開発と共有、及び公的教育と幅広い外部のステークホルダーの関与を通じて達成されるであろう[84]。AI に関するパートナーシップはまだ構築段階にあり、何も調査結果を発表してはいないものの、近い将来、議論に貴重な貢献をするのに必要な情報を有している。

英国王立協会[85]及び情報技術産業評議会[86]のような他の組織は、規格に関してさらなる作業が必要な領域を明示している。民間企業とコンソーシアムの作業に加え、12 を超える著名な研究機関が、多くの場合、AI 全般が生み出すものを確実に安全にすることを目的として、AI の発展が何を意味するかについて現在考察を行っている。

7.2 標準化のギャップ

この節では、前節で挙げた課題を解決するために対処する必要があるオープン標準化のギャップを幾つか挙げる。技術の発展とアプリケーションシナリオの増加に伴って、多くの規格を改善又は補足する必要がある。これらのギャップは、オープンソースと標準化の協調的な発展から、ニューラルネットワークの表現方法、パフォーマンス評価、機械学習のアルゴリズム、及びAI に関するセキュリティのギャップまでに渡る可能性がある。

7.2.1 調和的なデータモデルとセマンティクス

既に 6.2.2 節で強調したように、機械学習は、それを使って訓練がなされるデータに依存している。ただし、機械学習のアプリケーションは、様々な領域に分散している異なったデータソースに依存し得る。データがセマンティック能力を欠くと、それらは機械と人間の両方が理解できるやり方で処理されないため、AI システムはこれらの異なったデータソースを適切に使用することができない。さらに、データ

構造とフォーマットを統一する必要がある。特に機械学習の場合、多くのアプリケーションで、異なったソースからのデータが必要とされるため、この問題は重要な事柄である。

情報とデータモデルを標準化するための領域特有の取り組みは、既に現在存在している。しかし、これらの取り組みの間での調整が欠けている。AI の恩恵を得るには、異なる領域に渡る情報とデータモデルの標準化に対する同種のアプローチの調整を行い、確実なものにすることが必要であろう。

7.2.2 データモデルに基づく一般的オントロジー

異なった領域に、オントロジーに関する様々な作業が既に存在している。機械学習は一つの領域に限定されず、しばしば複数の領域を含むので、領域に依存したこれらのオントロジーを調和させる必要がある。異なる領域間で作業を調整し、一般的なオントロジーを標準化するための取り組みが必要とされる。

7.2.3 人工知能アルゴリズムの検証

AI アルゴリズムが、適用される全ての安全要求事項に確実に適合するようにするには、アルゴリズムの検証が必要である。AI アルゴリズムは実行中に変更される点で、主に他のアルゴリズムと異なっている。環境の変化も自己学習アルゴリズムの機能に影響を与える可能性がある。AI アルゴリズムのタスクを正確に文書化することは難しい課題であるが、このことは標準化によって恩恵を受ける可能性がある。どの要求事項に照らしてAI アルゴリズム又はそれが組み込まれているシステム全体を検証する必要があるのかを、明確にすることが望ましい。1 回だけの検証で有用なのか、又は周期的チェックが適切なのかも、問題になる点である[83]。

AIシステムは、AI製品の品質と利用可能性を確保し、AI産業の持続可能な発展を支援するために、機能、パフォーマンス、セキュリティ、コンプライアンス又は相互運用性など、様々な側面に関する認証の仕様を必要とする。認証は、試験、評価及びその他のタスクを含むことがある。評価の対象は、自動運転システム、サービスロボット、又はその他のAI製品が考え得る。評価の結果は、標準化された手順と方法に基づいた、測定可能な指標と定量化可能な評価システムを通じて得ることができる。

7.2.4 人工知能のインフラストラクチャのベンチマークと評価

革新的な設計、改良されたアーキテクチャ、及び新しいハードウェアコンポーネントを用いて、インフラストラクチャとプラットフォームの開発が常になされているので、機能、パフォーマンス、又はスケーラビリティに関してプラットフォームのベンチマークと評価を行うため、規格が必要とされる。インフラストラクチャの評価には、一般的な機械学習アルゴリズム又はシナリオ特有の作業負荷(例えば、条件の監視、監視ビデオ)のいずれかを使用することができる。ベンチマークと評価が、ユーザーが個々のニーズに最も適したプラットフォームを選択するのに役立つことがある。

第8節

結論及び推奨事項

この白書では、メガトレンドと呼ばれる最も急を要する課題の幾つかへの対処に役立つ幅広いアプリケーションに渡って、AIを採用できることの明確な証拠を示した。AIは既に社会、ビジネス及び個人に深い影響を与えているため、これは製造、エネルギー、金融、教育、輸送、ヘルスケア又は小売など、ほとんどの産業部門で根本的な革新の推進力の一つになることが予想される。

この白書では、四つのアプリケーション分野(スマートホーム、スマート製造、スマート輸送、及びスマートエネルギー)の具体的な使用事例を記載したが、複数の産業部門に渡っての現在のAIシステムの幅広い適用可能性も明確に実証されている。

社会、経済及び規制の分野でAIは深刻な課題を生み出している。信頼性、プライバシー、人/機械の相互作用の安全性、又は労働力に対する影響のいずれに関連するにしても、幅広いステークホルダーが相互に協調して、これらの問題に早急に対処する必要がある。これらの課題は世界の全ての国々に共通しているため、白書ではそれらに対処する国際的なアプローチが必要であるということを主張した。特に、政府、規制当局及び政策立案者は、AIの急速な採用と共に発生している主要な倫理的問題を、注意深く理解して対処する必要がある。

産業及び研究コミュニティは、複数のアプリケーション領域に渡ってAIの採用を阻害する可能性がある幾つかの技術的課題にも対処する必要がある。データ、アルゴリズム、ハードウェア及びコンピューティング・インフラストラクチャに関連する問題は、AIの潜在力を最大限に利用し、その最もエキサイティングな見込みを達成する上で、現在最も困難な制約の一つである。そうした観点から、新しいAI技術

が市場で採用され、社会に受容されることを促す上で、標準化と適合性評価が有益な役割を果たすことが期待される。

次の節では、産業全般、規制機関、そして最後に、IECとその委員会に対するこの白書の主な推奨事項の概要を示す。

加えて、さらに将来を見据えたアプリケーションを提示する。

8.1 産業界に対する推奨事項

幅広い産業界のステークホルダーが、AIエコシステムに関与している。こうした産業界のプレーヤーはAI技術から恩恵を受けることが予想されるが、彼らは質の高いデータの利用可能性、AIインフラストラクチャの継続的なパフォーマンスの向上、及びAIシステムのセキュリティの向上に対して、協調的な取り組みによって貢献する必要がある。これらの努力は最終的に、社会に対するAIの恩恵についての認識を高めることにつながり、AIの幅広い採用を加速させるであろう。

• データセットのガイドラインの作成

訓練データセットは、アルゴリズムのパフォーマンスにおいて最も重要な要素の一つである。アプリケーション領域に応じて、これらのデータセットには特別な要求事項が課される。訓練、妥当性確認及び試験のために、データ量がどのくらい多いことが望ましいかを推定するガイドラインを作成することが推奨される。これは企業がAIを良好かつ安全に使用することに役立つ。加えて、データの要求事項に関するガイドラインは、ビジネスが顧客のデータ

を求めるときに、ある中立的な存在を彼らに提供し、それを参照するようにすることで、ビジネスの助けとなるであろう。

- **AI システム内のセキュリティに関するガイドラインの開発**

AI は情報システムのセキュリティに深い影響を与える。一方で、これらは新しい対策に常に適応する新種のハックボット又はウイルスを作るために使用することができる。また、これは攻撃のための新しいアクセス点を与えることで、企業に対してさらなるセキュリティホールを作ることもできる。訓練データは、データセットの偏りを特に制御することにより、システムに影響を及ぼすことを試みるために使用できる。したがって、企業は適切なセキュリティ対策を講じるため、AI の適用前に、AI の使用がもたらすセキュリティ上の影響を明確化する必要がある。

8.2 規制に関する推奨事項

AI の採用には、規制面での大きな不確実性が伴い、このことは産業及び研究コミュニティから AI に対する投資の障害であるとみなされるであろう。また、AI が“欧州のデジタル単一市場(Digital Single Market in Europe)”のような、特定の政府プロジェクトに貢献していることは明らかである。したがって、市場への AI の導入を成功させるためには、複数のステークホルダー (産業、政府、研究コミュニティ、中小企業、消費者団体、その他)の間での協調的な規制措置が必要とされる。下記の規制措置が推奨される。

- **プライバシー方針の中で AI に対処する**

欧州の GDPR は、自動化された意思決定及びユーザーの同意がないプロファイリングなど、個人のプライバシー権侵害の側面の一部に対処している。しかし、対処する必要があるプライバシーに関する懸念はまだ幾つか残っている。

ハッカーが、個人からプライバシーデータを取り込むために訓練データをリバースエンジニアリングし、この機密データを悪用していることの証拠が存在している。また、アルゴリズムの設計と訓練のプロセスは網羅的である必要があり、不利な立場にあるグループの利益を十分に考慮することが望ましい。極端な倫理的状況に関しては、特別な規則を定めることが望ましい。この点について、セキュリティ、倫理及びプライバシーに関する関連規格、法律及び方針を早急に改善する必要がある。

AI がクラウド内に分布している場合には、個人が自分のプライベートデータをクラウドに送信するプロセスを十分に信頼でき、自らのプライバシーの侵害に関して心配がないことが重要である。さらに、AI の開発者は、データの収集と使用のプロセスにおいて個人を保護し、個人情報漏洩、改ざん及び破損を防止するために適切な技術的手段を採用する必要がある。

- **セキュリティ、倫理及びその他の方針に AI を含める**

現在の AI 産業は活発な発展の時期にあるが、将来の産業の監督手段の策定と実施について真剣に検討し、AI の開発と採用のプロセスにおける責任と欠点について十分に考察する必要がある。

AI 産業の監督の問題は、単一のグループだけが直面する問題ではない。これは、持続可能な AI エコシステムを構築するために、企業、政府、ユーザー、技術組織及びその他の第三者が参加し、共に作業することが求められる社会的、体系的で複雑な、幅広い問題を含んでいる。

関連する安全規制を策定し、改善することが望ましい。AI 技術の倫理的要求事項は、社会全体の中で形成された幅広い合意に従って設定されることが望ましい。

8.3 IEC とその委員会に対する推奨事項

世界的に認められた規格策定組織としての IEC は、今後の AI 技術開発の社会的な妥当性及び市場における妥当性を確保するために、この白書に記載した AI の課題の一部への対処を支援する独自の立場にある。IEC は、電気システムの安全性に対処するその実績と評判に基づいて、巨大な変化に対応し、AI がもたらす懸念を軽減する上で有益な役割を果たすことができる。そのために IEC は次の行動をとることが推奨される。

- 水平的な AI の標準化における JTC 1/SC 42 の中心的な役割を促進させる

JTC 1/SC 42 が作業計画を定めて標準化活動を開始したため、IEC はこの分科委員会を、世界規模での AI の基本的及び水平的な重要なエントリーポイントとして活かすことが望ましい。IEC はまた、ISO と密接に協力し、AI 技術に関連する側面について、IEC 及び ISO の個々の委員会の JTC 1/SC 42 との積極的な協力、及び考え得る共同作業を実施することが望ましい。作業の重複を避け、市場の迅速な受容を促すために、IEC は、標準化団体であれ産業コンソーシアムであれ、外部の組織に JTC 1/SC 42 との連携及び協力を勧めることが望ましい。

- データのセマンティクスとオントロジーの標準化を調整する

AI の成功は使用するデータの質と適切な解釈に大きく依存するので、データの種類、フォーム及び情報モデルを標準化することが、様々な産業部門での AI の受容と採用の成功のために重要である。セマンティックな相互運用性と標準化されたオントロジーは、AI 実装の一貫性と均質性を保証する中核的要素となる。IEC は、既存の作業と重複することなく、このデータ関連の課題に対処する既存の標準化活動の収束化の基盤整備、調整及び促進を、内部的に(例えば、SC 3D、領域特有の委員会である TC 57 及び TC 65 等、JTC 1/SC 32、JTC 1/SC 41)、及び外部の存在(例えば、ISO、eCl @ ss)に働きかけることによって主導することが望ましい。

- AI 関連の使用事例を展開し、一元管理する

AI はますます多くのアプリケーション領域に広がっていくので、IEC は、将来の IEC 使用事例管理場所 (UCMR) にさらに事例を含めるために IEC 62559 の方法論を適用することによって、AI 関連の使用事例を発展させることを、専門委員会に勧めることが望ましい。IEC は、使用事例の発展アプローチを推進するために、標準化団体であれ AI 標準化に関わる産業コンソーシアムであれ、外部の組織に働きかけていくことが望ましい。

- 一貫性のあるインタフェースをもつ AI 参照アーキテクチャを開発する

AI プラットフォームの開発を進めるにはには、AI システムの参照アーキテクチャを確立する必要がある。これにより、ソフトウェアとハードウェアの販売者は、自らの製品をエコシステムの中に位置付けて、コンポーネントとサブシステム間の相互運用性に貢献することができる。異種コンピューティング装置間のインタフェースと、AI タスクのスケジューリングとリソース利用のためのインタフェースは、異種コンピューティングプラットフォームにおける AI 作業負荷の配置をより簡単にするために、一致する必要がある。

- 潜在的 AI 適合性評価のニーズを探る

AI は、倫理と信頼性に関するかつてない課題と脅威を生み出しているため、AI アルゴリズムの適合性、及びそれらを利用する製品とサービスのアセスメントの必要性は大きいと予想される。ブラックボックス問題への対処、及び絶えず進化して自らの環境に適応するアルゴリズムの挙動の確認が非常に大きな課題になるであろう。IEC は、国際規格と適合性評価システムの両方を開発しているユニークな組織として、AI 適合性評価のニーズと機会を明確化し、ロードマップを作るための研究を開始することが望ましい。

- **社会の様々なステークホルダーとの AI に関する対話を広げる**

IEC は、関心をもつ全てのステークホルダー(産業、政府、規制当局、政策立案者など)を集め、社会と産業に対する AI の影響についての対話を開始し、広げていく上で理想的な立場にある。IEC はその国際的な地位と中立性を活かし、これらのステークホルダーのグループからインプットを集め、それを AI に関する自らの規格策定作業へ含め、そのプロセスへの信頼を継続的に構築し、その成果が AI 関連の規制と立法作業に確実に使われるようにすることが望ましい。

- **IEC が関与するテストベッドに AI の使用事例を含める**

エッジインテリジェンスに関する以前の IEC 白書 [12]は、IEC の標準化活動に関してフィードバックを得るためにテストベッドを採用することを推奨した。この推奨の意図に沿うなら、テストベッド内の AI 使用事例は、AI 技術のギャップ及び標準化ニーズに関してフィードバックを収集する上で不可欠である。AI 用に特定のテストベッドを導入することの妥当性を示すことは難しいが、標準化と技術ギャップについて明確化するには、AI 関連のアーキテクチャと規格主導の AI 実装に関するフィードバックが不可欠である。さらに、新しい機械学習の機能、アルゴリズム、データ及びオントロジーモデルは、AI 使用事例を含んだテストベッドから恩恵を受けるかもしれない。したがって、IEC に関与する将来のテストベッドに、AI 関連の使用事例を加えることが推奨される。

附属書 A

将来の発展

A.1 生物学に触発された人工知能

ニューラルネットワークは、人間の脳がどのように機能しているかに触発されているが、それを模倣しようとはしていない。現在、脳が実際にどのように機能しているかはまだ完全には理解されていないので、実際には、人間の脳のメカニズムをコピーすることは非常に難しいように見える。しかし、人間の脳で学習がいかになされるかをより良く理解することが、優れた機械学習アルゴリズムにつながるという点に、おおむね意見の一致が存在している。これは汎用的な AI を作る上で不可欠なステップであり得ると考える人もいる。

脳内で学習メカニズムがどのように機能しているかについては、幾つかの理論があり、これらの理論に基づいた幾つかのアルゴリズムが開発されている。それらは、脳の正確なコピーを作ろうとはおらず、通常は新皮質に限定された、脳の中央領域の一部を模倣しようとする。大脳辺縁系など、脳の他の領域は、これまで機械学習コミュニティからあまり注目されていない。これらのアプローチの一つにより人間の脳を何らかの形で再現できるかどうかはまだ明確でないが、それでも近年、幾つかの注目すべき開発がなされている。

これらのアプローチの一つは、階層的一時記憶 (HTM) である。一見、この構造はニューラルネットワークのように見えるが、それとは根本的に異なっている。HTM は階層的に構造化されたノードのネットワークである。これらのノードは情報を学習し、保存する能力をもっている。この構造は、時間変化するデータを階層的に表示するように設計されている。HTM 内のデータのコーディングは、疎分散表現 [114] と呼ばれる形式で利用でき、これはパターン内のビ

ットベクトルから構成されていて、そこでは各ビットがセマンティックな意味をもつ。このモデルは、人間の脳がどのように学習するかについての理論から触発されたものであった [115]。

結局のところ、現時点では、近い将来に汎用 AI が作られる可能性はまだないように見える。しかし、人間の脳に関する欧州プロジェクト (European Human Brain Project) [116] のような脳研究における最新の発見は、新しい種類の機械学習アルゴリズムに勢いをもたらし得る。

A.2 人間/人工知能の相互作用

AI の使用は、人間-機械相互作用の新しい機会を開く。これまで機械は通常、特定の方法で操作される単なるツールであったが、洗練された AI アルゴリズムは将来、機械をより良く人間に適応させることができる。

こうした改善された相互作用の初期段階は、モバイル装置又は家庭でしばしば使用されている多くの支援システムにおいて既に見られる。しかし、現時点では、機械に適応する必要があるのは一般に人間の側である。機械が処理できるような特定のやり方でリクエストを構成する必要がある。ただし、最終的な目標は、機械が人間に適応することである。結局のところ、AI は人間-機械相互作用が円滑に機能できるようにするインタフェースの役割を果たす。

こうした相互作用の一例は、電話サービスのホットラインである。ほとんどの自動留守番機械は、発信者が入力した特定の単語又は番号にだけ応答する。こうした自動クエリは将来、発信者の質問及びニーズに動的に応答する AI システムによって実行される

ので、フラストレーションとストレスを回避することができる。

AI は、新しい安全コンセプトの中心的要素の一つであり、とりわけハンドリングマシンに関してそうである。これは、特にロボティクスにおいて重要である。現在、家庭で見られる最も自律的なロボットは掃除機ロボットである。しかし、将来、はるかに大型の、おそらく人間型のロボットが、家庭又は公共環境で使用される可能性がある。後者の場合に安全性を確保するため、そうしたロボットは自らの環境、特にその中にいる人間を認識できなければならない。変化する周囲環境についてのこうした先進的な理解は、先進的な機械学習技術によってのみ可能である。

A.3 人工知能利用デジタルツイン

デジタルツインは物理的な対象物とプロセスのデジタルモデルであり、これは、接続されたデバイスが自らとその周辺に関するデータを収集して共有するIoTによって可能となる。物理的なデバイスが収集したデータをリアルタイムで結合することで、デジタルツインは、自身の物理的な対応物の状態を模倣する。その後、こうしたバーチャルなツインを、物理的な対応物のパフォーマンスを向上させるために分析し、実験することができる[117]。したがって、デジタルツインの登場は、エンジニアが新しい設計をバーチャル空間で試験できるようにするだけでなく、それらのパフォーマンスの継続的な監視が可能になることで、製品開発に深い影響を与える。その結果、AI はそうしたシミュレーションを支援し、デバイスの将来のパフォーマンスを予測し、考え得る改善について推奨を行うことができる。デジタルツインの使用の興味深い適用は、例えば、性能をより良く理解し監視するためのタービンへの適用、又は遠隔地にある機器への適用を含んでいる。実際に設置する前に、問題をより容易に発見する、又はシナリオの試験を行うことができ、それによって新しい又は通常と異なる場所の様々な環境要因を明らかにすることができる。

特に、ビデオやオーディオのような、数値センサによる読みで対応できないようなデータストリームを含む場合に、AI はそれらを分析に含めることで大きな価値を付加することができる。発生しかけている機能不全の兆候を検出し、事前対応的保守を可能にするために、オーディオフィードを分析することができ、欠陥を非常に正確かつ迅速に明確化するために、生産プロセスのリアルタイムビデオを使用できる[118]。

AI は、この種のデータの高速度での分析を可能にする以外に、AI を利用したデジタルツインと人間の間の生産的な相互作用を促進することができる。AI を利用したロボットを、人間とデジタルツイン(あるいは、さらにいえば、何らかの物理的な対象物)の間の相互作用を向上させるために活用でき、それによってユーザーエクスペリエンスと生産性が改善される。ユーザーは、労力を費やして技術的なログを調べたり、AI システムから単にコマンドを受け取ったりするのではなく、与えられた機械的コンポーネントのデジタルツインとの間の相互作用に、自然言語を使用して、彼らが人間の同僚とそうするのと同様に、AI システムと一緒に問題と機会について検討することができる[119]。

この開発は、人間の生産性を高めるだけでなく、その機能が人間の労働に置き換わるよりはむしろ、その支援である点を強調することによって、こうしたシステムの受け入れを大きく増加させるであろう。実際、フルに実装されデジタルツイン機能を用いて、組織は人的資本と機械的資本の両方を含む労働力を創出することができる。

A.4 自動化された機械学習

今の時点で、自動化された機械学習の様々な例が存在している。一つは強化学習であり、これは生の入力データから直接学習するためにエージェントを使用する。エージェントは自らの環境と相互作用し、自らの行動に応じて報酬又は罰を受ける。報酬の最大化を追求することによって、エージェントは正しい選択をすることを学ぶ。例えば、ゲームのルール

について追加の知識を得ることなく、50 のアタリゲームをいかにプレーするかを学習することに成功した強化学習アルゴリズムが存在する[120]。

自動化された機械学習への別のアプローチは、機械学習モデルの開発を自動化することである。こうしたシステムは、自律的に新しい子ネットワークを作り、それらを訓練する。次に、システムは自らのパフォーマンスに応じて、その後の試行のための改善点を見出す。このようにしてモデルはネットワークを改善していき、アーキテクチャのどの部分が良い結果を達成するかを学習する[121][122]。

自動化された機械学習を訓練とモデル化の労力を減らすために使用する傾向は、将来さらに加速する可能性が高い。これは、開発したアルゴリズムをいかに制御下に置き続けるのか、及び様々な AI アプリケーション間の対話がますます増えて互いを改善していく可能性がある事実がいかに対処するかという問いをもたらす。

1010 1000 0100 0110
0100 1010 1010 1010
0111 0100 0111 1010 1000 1010
1000 0100 1010 0110 0110
1010 1010 1000 0111 0100 1010
0111 0100 1010 1000 1010

1110 1010 0110 0010 1010 0010 0111 0010 0110
0110 0110 0010 0101 1010 0110 0111 1010 0010 1110
0110 0010 1010 0010 0111 0101 1110 1010
110 0111 1010 0010 0110
0110 0111 0010 0010 0110 0101 1010
0010 0111 0110 0010 0010 0110
0010 1010 0101 1010 0111 1110



International
Electrotechnical
Commission

ISBN 978-2-8322-5874-3



CHF 50.-

3 rue de Varembé
PO Box 131
CH-1211 Geneva 20
Switzerland

T +41 22 919 0211
info@iec.ch
www.iec.ch

© Registered trademark of the International Electrotechnical Commission. Copyright © IEC, Geneva, Switzerland 2018