

White Paper

AI 実装が広がる世界のアプリケーション開発

Sponsored by: Amazon Web Services

Ritu Jyoti
April 2019

David Schubmehl

調査概要

熾烈な競争が繰り広げられる現代のビジネス環境の中で、各企業は顧客や株主、従業員に対してより高い価値を生み出し提供する方法を求め続けている。特に、さまざまなタイプの自動処理と増大し続けるデータとを組み合わせることで売上拡大、コスト削減、ビジネスプロセスの合理化、そして顧客をより深く理解するための新たな方法を追求している。その中で、自社のアプリケーションやビジネスプロセスに実世界の AI（Artificial Intelligence：人工知能）を組み込む手段として機械学習や深層学習（深層学習）に目を向ける企業が増加している。機械学習とは、各種データ群を基に、人がプログラムすることなくさまざまな機能を実行できる統計モデルを構築するプロセスである。深層学習とは機械学習の一種で、ニューラルネットワークのアルゴリズムに基づき大量のデータを使用したトレーニングを経て、より正確な分析、提案、予測を得るために用いられる。

各企業は深層学習モデルを、製品提案、価格予測、画像認識、意思決定の改善など多様な目的に活用している。深層学習モデルの開発には最近まで相当な時間、労力、知識、専門的知見が必要とされていた。しかし、AWS（Amazon Web Services）などのベンダーが深層学習用のサービスやツールを開発したことで、データサイエンティストや開発者が深層学習モデルを簡単かつ迅速に開発、導入できるようになった。

機械学習や深層学習向けのツールやフレームワークは、TensorFlow、Apache MXNet、PyTorch など数多くに提供されており、いずれもインテリジェント型アプリケーションの開発に有効な機能を備えている。ただし、機械学習アプリケーションの開発にはさまざまな条件が存在し、対応を間違えればむしろ足枷となってしまう。

- 現在の作業に適切な機械学習、深層学習用のフレームワークを選ぶ
- 適切な機械学習、深層学習のアルゴリズムを選ぶ
- 最も正確な予測が得られるように、機械アルゴリズムとデータを調整、チューニングする
- 機械学習モデルのトレーニングに使うデータを確定し、処理の場所を決め、収集する
- モデルのトレーニングと本番での予測生成（推論）のどちらにも対応できる適切なコンピューティング処理能力を確保する
- 機械学習モデルや深層学習モデルを既存の業務アプリケーションに統合する
- 本番の大規模運用時に十分な性能を引き出せるようモデルの運用体制を整備する

クラウドをベースとする Amazon SageMaker（機械学習向け管理プラットフォーム）や AWS Deep Learning AMI（事前に構築された実行環境、AMI は Amazon Machine Images の略）は、こうした機能の多くを提供しており、機械学習や深層学習を取り入れたアプリケーションをいち早く市場に投入したい開発者とその企業にとって有効なツールである。

すでにトレーニング済みの深層学習モデルを簡単、便利に使いたいと考える企業は、次に列記するような API（Application Programming Interface）を利用することで、インテリジェント機能を短時間でアプリケーションに追加できる。具体的には、画像や動画用の Amazon Rekognition、チャットボット統合用の Amazon Lex、テキスト読み上げ用の Amazon Polly、自然言語翻訳用には Amazon Translate、音声認識用には Amazon Transcribe、顧客ごとにカスタマイズされたレコメンデーション用には Amazon Personalize、時系列の正確な予測には Amazon Forecast、ほぼあらゆるドキュメントからテキストやデータを抽出できる Amazon Textract、構造化されていない臨床テキストから有益な情報を発見する Amazon Comprehend Medical、テキスト内の相関性を導き出す Amazon Comprehend などの API が用意されている。開発者にとっての利点は、カスタマイズした機械学習モデルをすべて自分で構築する代わりに、こうした API をそのまま単純かつ容易に利用できることである。ほとんどの開発者にとって、機械学習モデルの独自開発に取り掛かる前に、すでに用意されている API を使うことで自分たちの問題が解決できる可能性を確認することが有益であろう。

概況

はじめに

機械学習や深層学習を利用した AI アプリケーション市場は急成長を続けている。機械学習／深層学習ソリューションへの支出は 2021 年には 570 億ドルに達すると IDC では予測している。2026 年にはエンタープライズ向けソフトウェア全体の 75% には何らかの形で機械学習あるいは深層学習による予測、提案、助言機能が組み込まれると IDC ではみている。こうしたシステムがなぜ各企業の将来にとって重要であるかについて、各企業は以下の項目について十分に検討する必要がある。

- **人間の判断力を高める**：最良のビジネスケースは、AI 対応アプリケーションを人間の意図の延長線上に位置付け、人間の能力を置き換えるのではなくむしろ拡張する用例である。職人の使う電動工具が最も分かりやすい例である。この分野での好例として、価格設定最適化モデルが挙げられる。あるいは、放射線専門家による X 線画像解析をサポートする AI 画像アプリケーションも適切な例である。
- **調査、発見に必要な時間を短縮する**：最強の読解力を持つ読み手でも 1 日に数百万ページは読了できない。代わりに自然言語を理解できるアプリケーションを使えば、音声にも文字にも対応できる。深層学習型自然言語システムは、分類をベースに人間が作り上げたシステムよりも優れた成績を示している。
- **「次に打つべき最善の手」を提案し結果を予測する**：最も一般的なユースケースとして、相関性の高いデータに基づき提案や予測を行う深層学習型アプリケーションなどが考えられる。
- **結果と提案をパーソナライズする**：特定の顧客や見込み客に合わせたコンテンツや予測、提案の内容を「パーソナライズ」する目的で深層学習モデルを活用する企業が増えている。自分のデバイスやアプリケーションが、その利用者である自分の好き嫌いや期待について「よく知っている」であろうモバイルアプリケーションでは、これはまさに正解と言える。

- **組織のナレッジを自動管理する**：数十年前にナレッジ管理システムが登場して以来、運用にかかる労力がネックとなり多くの製品が消えていった。この領域では調査や発見、ベストプラクティス策定の自動化こそが主なベネフィットとなる。ドキュメントの分類やテーマ決定などは深層学習のポイントとなるユースケースである。他のユースケースとしては、新しい分析レポートを作成、新たなプロジェクト用コンテンツを開発するナレッジワーカーに関連するコンテンツやヒントを提示できるオフィス用アプリケーションが考えられる。
- **ベストプラクティスをパッケージ化、「システム化」する**：この分野では、ルール型やヒューリスティック型のシステムに代わって、前述した経験に基づくテーマ決定を応用した機械学習モデルの構築が主なユースケースとなる。

各企業は、深層学習を実装したアプリケーションを用いることで、ビジネスプロセスの破壊と再構築、デジタルトランスフォーメーション（DX）、新たなスケールメリットの創出に取り組んでいると IDC では分析している。大手ヘルスケア企業は、クライアントである患者がどこで処置を受けるかに関係なく、公平で迅速な診断と治療の「ベストプラクティス」を実現するために、クラウドベースの深層学習によるコンピュータビジョン（画像認識）アプリケーションがどのように寄与できるかについて研究している。グローバル金融企業では、AI 金融取引のワークフローやビジネスプロセスに残された手作業部分について、アプリケーションによる時間短縮、自動化、あるいは廃止を進めている。メーカーは IoT や深層学習モデルに基づく予防保全戦略を編み出しつつある。その他にも何百というユースケースが存在し、市場や競合他社が AI、機械学習、深層学習を活用したアプリケーションを採用し始める状況の中で、各社が研究に取り組んでいる。

AI を実装したアプリケーションの台頭に伴い、深層学習に向けたオープンソースに基づくツールやソリューションの市場が拡大している。意欲的で有能な開発者、実績あるオープンソースのコミュニティ型開発モデル、コンピューティングシステムの学習用に低額または無料のソフトウェア製品を利用したいという要望、この 3 つの強力な組み合わせによって、機械学習や深層学習のソフトウェア向けにライブラリーやツールを作成する市場が形成されつつある。こうしたツールは、音声認識、自然言語処理、予測分析、アドバンスドアナリティクス、ニューラルネットワーク、教師あり、または教師なし機械学習などに及ぶ広範囲なテクノロジー群の一角を占めている。データを与えた後は自力で学習しプログラムを作成してさまざまな問題を解決できるように、自己学習アルゴリズムと特殊な組み込みライブラリーによってアプリケーションをスマート化するのが最終目標である。こうした自己プログラム型コンピューターアルゴリズムこそが、IDC がインテリジェントアプリケーションと呼ぶ製品興隆の原動力である。

機械学習や深層学習向けサービスとしてのツールやフレームワーク、ライブラリーの勃興によって、今日では開発者が低コストでインテリジェントアプリケーションを構築できるようになった。各企業では、プログラミングやメンテナンスに膨大な手間のかかるルール型やヒューリスティック型の手法を離れ、こうした新サービスを採用する動きが見られる。ハイパフォーマンスコンピューティングリソース、膨大なデータ、機械学習と深層学習用のフレームワークやライブラリーは、プログラミングなしでの課題解決を可能とした。こうした機械学習と深層学習用のライブラリーやテクノロジーのユースケースはさらに広がり、画像認識や病理診断、価格設定最適化や商品レコメンデーションに及んでいる。機械学習と深層学習は大部分の AI アプリケーションの根幹となるばかりでなく、エンタープライズ向けアプリケーションでも採用が始まっている。機械学習のバリエーション、効率性、信頼性が向上するにつれ、こうしたアプリケーションの有用性と安定性が改善され、普及が進むと考えられる。

深層学習

深層学習とはニューラルネットワークアルゴリズムに基づく機械学習の一種であり、大きな商業的成功を収めている。ニューラルネットワークは、アルゴリズムが適切な回答に到達するように、人間の脳に倣って相互結合された人工ニューロンに割り当てられた重みを調整する。深層学習モデルは、インサイト、助言、予測の精度をより高いレベルで生成できるように、時間をかけて各ニューロンの重みの調整を行い（このプロセスをトレーニングと呼ぶ）、重みに影響を及ぼすテキスト、音、その他のデータ、の複雑なパターン認識の反復を通じて改善を進める。

深層学習には多数のフレームワークやツールが存在する。以下にその一部を示す。

- **TensorFlow**：グーグルが開発した、オープンソースソフトウェアに基づく数値計算用フレームワーク。
- **Apache MXNet**：分散（ディープ）機械学習コミュニティ（Distributed (Deep) Machine Learning Community、DMLC）が構築した、スケーラブルなトレーニングと推論用のフレームワーク。Apache ソフトウェア財団のインキュベーションプロジェクト。
- **PyTorch**：「GPU ファースト」を掲げる N 次元配列対応深層学習用フレームワーク。各種配列データ処理用の関数ルーチンライブラリーが提供されている。PyTorch は Torch を利用した Python パッケージで、高度で複雑な AI 処理の開発と機械学習に適している。

こうしたフレームワークの他、Gluon や Keras など、より高いレベルで用いられるツールもある。Gluon は深層学習用 API で、初心者から上級者までの深層学習モデル開発者にプロトタイプ、ビルド、トレーニングの機能を提供している。現在 Gluon は深層学習用フレームワークの Apache MXNet に対応し、開発者が簡潔なプログラム用フレームワークを通じニューラルネットワーク型モデルを短時間で開発、トレーニングできるように考慮されている。さらに、Gluon は、モデルトレーニングの効果を段階的に評価することでニューラルネットワーク変更の容易さとトレーニングモデルのデバッグ迅速化を実現、開発者のアジャイル性を高めている。Keras は高レベルなニューラルネットワーク API で、Python で記述され TensorFlow や Microsoft Cognitive Toolkit 上で作動する。モデル開発時間の短縮を主眼に開発された。

深層学習の意味

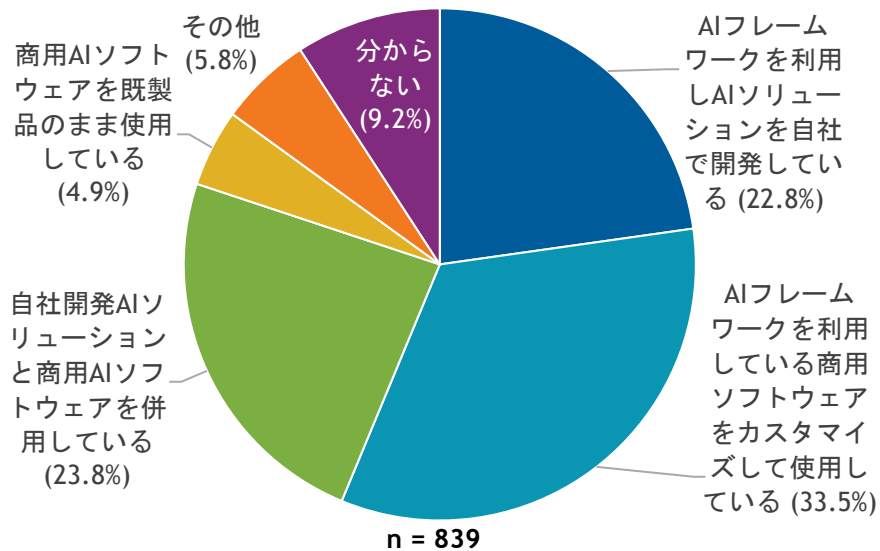
IDC 調査によると、85%の企業が AI ソリューションを積極的に開発している（Figure 1 を参照）。今回の調査におけるコグニティブ/AI システムの定義は以下の通りである。

自然言語処理、音声認識、機械学習、ナレッジグラフその他の技術を利用して質問に答え、知見を見出し、提案を行う一連の技術。このようなシステムは与えられた事実を基に仮説を構築し何らかの解答を提示できるほか、大量のコンテンツの取得によりトレーニングが可能であり、成功や失敗を通じて自動的に補正、学習する。

FIGURE 1

コグニティブ/AI システムの導入状況

Q. あなたの組織における AI ソリューションの開発と利用の状況について教えてください。



n = 839

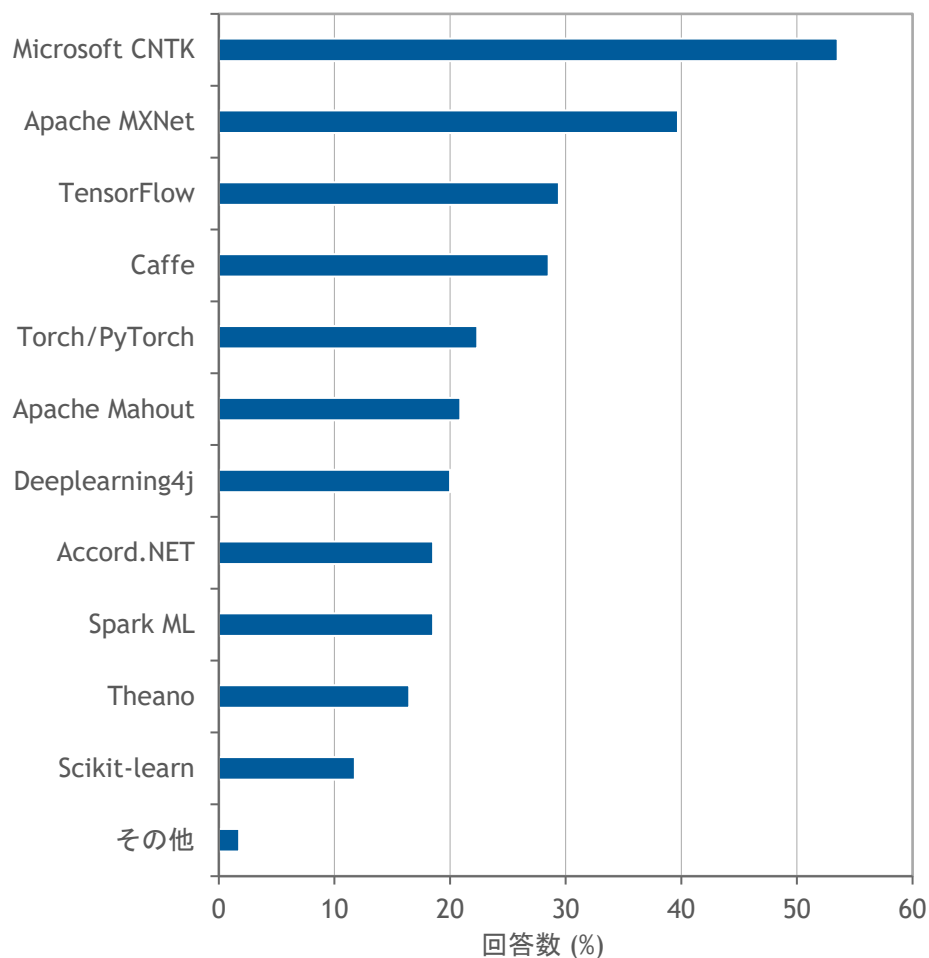
Source: IDC's AI Software Platform Adoption Survey, February 2019

さらに、AIソリューションを自社開発、あるいは商用ソフトウェアをカスタマイズして使っている組織は、いずれもさまざまな機械学習フレームワークも利用している（Figure 2を参照）。

FIGURE 2

機械学習フレームワークの導入状況

Q. 使用している機械学習フレームワークの種類は何ですか？



n = 340

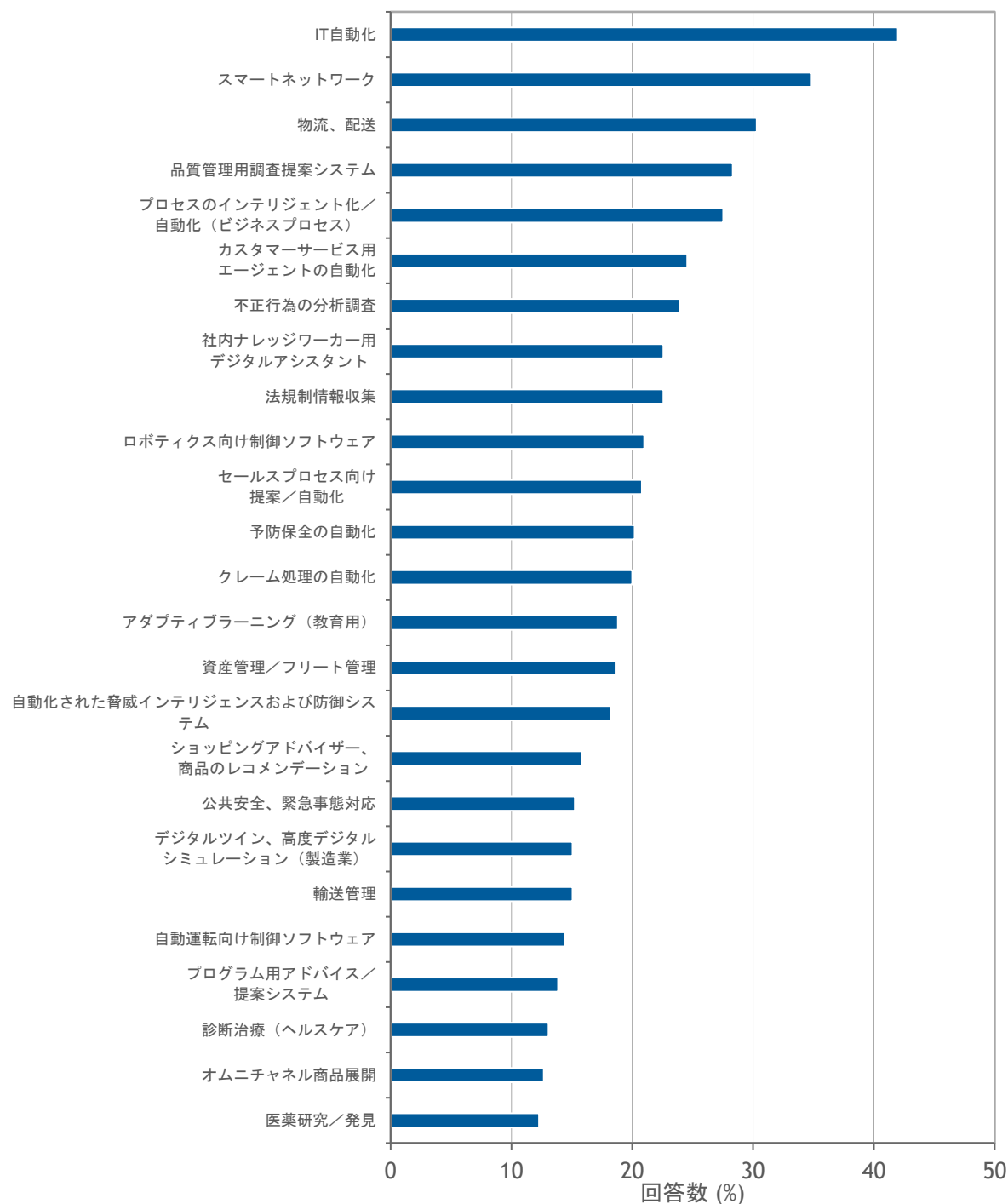
Source: IDC's AI Software Platform Adoption Survey, February 2019

調査対象となった組織のほぼすべてが、広範なユースケースを念頭に AI アプリケーションを開発していると回答している（Figure 3を参照）。

FIGURE 3

AI アプリケーション

Q. 開発したもしくは開発中のAI アプリケーションの対象となるユースケースは何ですか？



n = 505

Source: IDC's AI Software Platform Adoption Survey, February 2019

機械学習と深層学習を採用した AI アプリケーション開発は、エンタープライズ向けソフトウェア市場とアプリケーション開発の両市場で人気が高まっている。とはいえ、こうした新技術の採用を目指す企業は大きな課題に直面している。

- 異なる種類のツールと技術が多数存在するものの、どこから手を付け、どのようなツールや技術を用いるべきか見極めるのは企業にとっても開発者にとっても難しい。
- 機械学習用の統合開発環境が存在しないため、実験、開発、テスト、本番導入のサイクルを迅速に回せない。各開発言語用ツールは以前から存在するものの、モデルを本番環境に導入する手順が複雑で時間もかかっていた。ただし、現在ではツールの改良によって改善されている（AWS の詳細な例は後述する）。
- 機械学習やデータサイエンスに関するスキルが開発者側に十分に備わっていないことが、各組織にとっての問題を大きくしている。データサイエンティストの採用は容易ではない。既存の機械学習フレームワークを活用できるスキルを持った開発者を見出すのはさらに難しい。
- 事前構築済みドメインに用いる API やテンプレート式ソリューションも存在するとは言え、それを見付けてプロジェクトに利用するのは容易ではない。

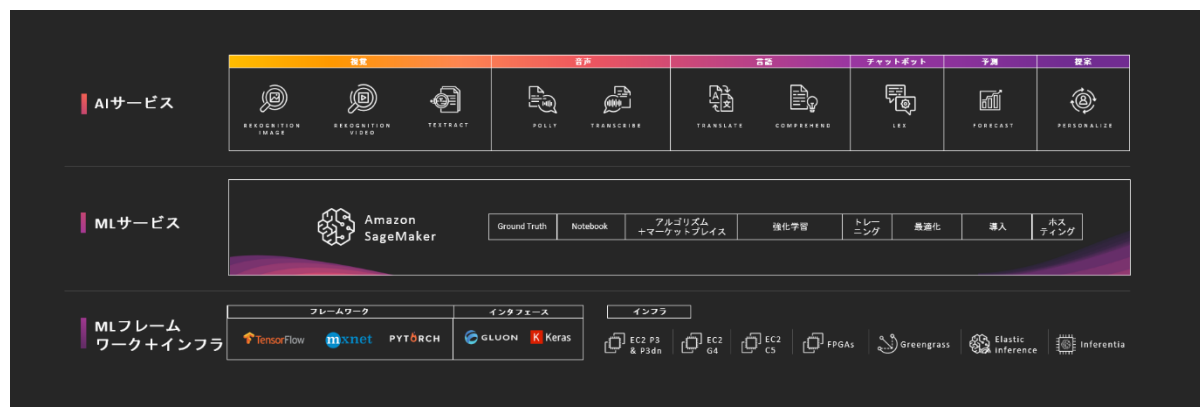
最小の時間と労力で高精度な予測型あるいは処方型モデルを構築するのに最適な方法を見出すことが、企業にとっての課題である。

AWS の機械学習、深層学習用ツール

AWS が提供する広範なツールやサービスによって、深層学習モデルの構築と本番導入に費やす機械学習や深層学習関係作業は、開発者にとって極めて容易なものになった。Amazon Elastic Compute Cloud (EC2) の P3 インスタンス経由でインフラとして簡単に使える NVIDIA Volta を活用した GPU サービスを始め、TensorFlow や Apache MXNet などフレームワーク、Spark や Amazon EMR などプラットフォーム、新発表の Amazon SageMaker など各種ツールが存在する。加えて、Amazon Comprehend、Amazon Transcribe、Amazon Rekognition など深層学習型サービスも提供されている。各サービスのアーキテクチャを Figure 4 に示す。AWS はすべての深層学習用フレームワークに対応しており、この点でいずれかのフレームワークを推奨する他のクラウドサービスプロバイダーとは異なるアプローチをとっている。

FIGURE 4

AWS 機械学習アーキテクチャ



Source: AWS, 2019

Amazon SageMaker : 機械学習用のフルマネージドサービス

機械学習モデルの開発、トレーニング、導入を容易にするために登場したのが、AWS の Amazon SageMaker である。SageMaker はフルマネージドサービスで、すぐに機械学習を利用したいという開発者を悩ませていた時間のかかる問題をすべて取り除き、あらゆる規模の機械学習モデルの構築、トレーニング、導入に関わる開発者やデータサイエンティストの作業を簡素化、迅速化している。開発者の足枷となっていた各工程の複雑性を取り払うのが SageMaker の役割である。SageMaker は、機械学習モデルの構築、トレーニング、導入に対し、単独、併用いずれでも使用可能なモジュールで構成されている。

まず、機械学習を使った高精度なトレーニングデータを短時間で構築し、データのラベリングにかかるコストを最大 70%削減できる Amazon SageMaker Ground Truth がユーザーをサポートする。機械学習モデルの成功には、ラベル付きデータを使ってモデルに正しい決定方法を教えるトレーニングが欠かせない。このプロセスには大規模なチームでも数か月かかるのが一般的であったが、機械学習と人手によるラベル付けを統合した SageMaker Ground Truth の革新的ソリューションによって、コスト削減や複雑性排除と、データラベルの高精度化を両立させることが可能となった。これにはアクティブラーニングと呼ばれる技術が活用されている。

設定作業が不要である SageMaker は Jupyter Notebook をホスティングしているため、開発者はトレーニング用データセットの処理や機械学習、深層学習用モデルの開発をすぐにスタートできる。SageMaker コンソールを数回クリックすればフルマネージドな Notebook 用ワークスペースが完成する。自社 VPC へのセキュアなネットワーク接続を確立し Amazon EC2 インスタンスを立ち上げるところまで対応するサービスには、TensorFlow や Apache MXNet など機械学習や深層学習用のフレームワークに使えるライブラリーがあらかじめ実装されている。開発者による Notebook の自作やインポート、あるいはリスクモデリングやチャーン（顧客の離脱）予測または OCR など一般的ユースケースに基づく既成 Notebook へのデータ取り込みにも対応している。Amazon S3、Amazon Redshift、Amazon DynamoDB、Amazon RDS などのモデルトレーニング用データを揃える際には、AWS Glue や Amazon EMR 上の Apache Spark でデータを事前処理するほか、ワークスペース用の追加ストレージとして Amazon EFS を利用できる。

アプリケーションのトレーニングを開始する準備が整えば、あとは開発者が、必要な Amazon EC2 インスタンスの種類と数を指定しクリックするだけでトレーニングが始まる。SageMaker が分散コンピューティングクラスターを構成、トレーニングを実行し、作業完了時点でクラスターは解体されるため、ユーザーに課金されるのは実際に使ったリソース分のみであり、基盤となるインフラを意識する必要はない。SageMaker は事実上無限大のノード数までシームレスに拡張可能なため、開発者は分散型トレーニング用アーキテクチャの構築、運用に伴う煩わしさから解放される。

SageMaker はスピードとスケール、精度に照準を合わせた高性能でスケーラブルな機械学習アルゴリズムを提供する。正解が用意された教師付きアルゴリズムを選択してトレーニングを実施し、モデルが誤った箇所を指摘できる。SageMaker には XGBoost を始め、線形回帰やロジスティクス回帰あるいは類別など教師付きアルゴリズムが実装されており、提案や時系列予測問題に適用可能である。SageMaker には k 平均法や主成分分析（PCA）など教師なし学習（すなわちアルゴリズムが自力で正解を探し出す）機能も備わっており、購入行動から顧客グループを割り出す場合に有効である。

トレーニング用データが存在しない場合は、強化学習に対応した Amazon SageMaker RL を使って機械学習モデルに決定方法を教える。Simulink や MATLAB などシミュレーション環境や MXNet、TensorFlow、Intel RL Coach などフレームワークに対応した SageMaker RL は、広範なユースケースに適用可能である。

SageMaker を使えば、深層学習モデルのチューニングに必要な時間も短縮できる。モデルが可能な限り高精度になるように数千通りものアルゴリズムパラメーターの組み合わせを自動調整するため、数日がかかりで試行錯誤しながら人手で調整する手間が不要となる。

Amazon SageMaker Neo を使えば、一度トレーニングした機械学習モデルはクラウドやエッジのどのような場所でも実行できるようになる。複数のプラットフォームで実行できるよう機械学習モデルを最適化するためには、プラットフォームごとに個々のハードウェアやソフトウェアの構成に合わせて開発者が手作業でモデルをチューニングしなければならない、一般的には大変困難である。Neo は ARM やインテル、NVIDIA、Cadence、クアルコム、Xilinx の各プラットフォームに合わせて TensorFlow や MXNet、PyTorch、ONNX、XGBoost の各モデルを自動最適化することで、この時間と労力を不要とした。Neo によって、クラウド上のインスタンスやエッジデバイスに最適化された機械学習モデルが数クリックで構築可能となった。従来型モデルと比較した場合、最適化済みモデルのスピードは最大 2 倍、かつストレージ容積は 10 分の 1 以下である。

SageMaker は、トレーニングの成果物であるモデルアーティファクトやスコアリングイメージを提供するため、開発者はそれを見て EC2 など任意の環境に導入する。次に、開発者が EC2 インスタンスの種類と数を指定した後、SageMaker がユーザー組織のアプリケーションに合わせてインスタンス起動、モデル展開、HTTPS エンドポイント設定を実行し、低レイテンシで高スループットのインファレンスが実現される。本番環境への導入後は、SageMaker がコンピューティング用インフラを管理してシステム診断やセキュリティパッチ適用などの定常メンテナンスを実施する。Amazon CloudWatch が組み込まれているため監視とログは常に有効である。Jupyter Notebook のホスティング、モデルのトレーニング、予測実施、出力ログ作成に使用した AWS コンピューティングやストレージのリソースが課金対象となる。構築、トレーニング、ホスティングのコストは秒単位で課金され、基本料金も前払いも必要ない。

新モデルの統合時も SageMaker であればコード更新不要である。Amazon SageMaker には A/B テスト機能も実装されており、モデルテストや別バージョン試験によって最善の結果を得たいユーザーをサポートする。

Amazon SageMaker で Inference Pipelines を導入した場合、ローデータの取り込みと事前処理、予測、事後処理に関するリアルタイムやバッチによるインファレンス要求に対応可能となる。Amazon SageMaker で利用可能な機械学習フレームワーク、組み込みアルゴリズム、カスタムコンテナであれば、任意のコンポーネントを使って Inference Pipelines を構成できる。Spark ML が提供する独自のトランスフォーマースイートや Amazon SageMaker が提供する Scikit 学習フレームワークコンテナによって独自のデータ処理やエンジニアリングに用いるパイプラインを構築し、Inference Pipelines の一部として展開することで、データ処理コードの再利用と機械学習プロセス管理の簡素化が実現する。

最後になるが、モジュラー型アーキテクチャは SageMaker の最も優れた特徴と言える。開発者は組織のワークフローに対応して構築、トレーニング、ホスティングの機能を自由に組み合わせていく。SageMaker を活用することで、Notebook 用テンプレートとアルゴリズムを選択した後に管理サービスを使ってトレーニング、テスト、導入を実施するだけで機械学習モデルを開発、導入できる。すなわち、SageMaker が提供するエンドトゥエンド型機械学習環境は、モデルを構築、トレーニングし、本番アプリケーションとして投入するまでの工程を大幅に加速、簡素化する力を持っている。

AWS 深層学習 AMI

Amazon は企業と開発者向けに AWS Deep Learning AMI も提供している。EC2 上に構築された AWS Deep Learning AMI には、深層学習アプリケーションの構築、トレーニング、実行に必要な環境が完備されている。AMI には Apache MXNet や Gluon、TensorFlow、Microsoft Cognitive Toolkit、Caffe2、Torch、PyTorch、Keras などオープンソースの深層学習用フレームワークがあらかじめ実装されている。いずれも GPU アクセラレーション用に CUDA と cuDNN の各ドライバーが設定済みである。その他にも人気の高い Python パッケージや Anaconda Platform が組み込まれている。プラットフォーム上に各インスタンス用に Jupyter Notebook も設定されており、Python 2.7 または 3.4 を使ったインタラクティブな深層学習モデルの導入に対応する。AWS Deep Learning AMI には、Jupyter 以外にも CppLit や PyLint、Pandas、GraphViz などのツールキットが用意されている。

AWS Deep Learning AMI には 3 つのバージョンが存在する。一つが Conda 準拠 AMI であり、オープンソースパッケージの環境管理ツールとして人気の高い Conda で構築された深層学習用フレームワークが個別の Python 環境として準備されている。2 点目が Base AMI であり、開発者自身がカスタマイズした深層学習モデルを展開できるよう GPU ドライバーとライブラリーを備えている。こうした設定済み環境の存在によって、開発者は自由かつ柔軟にツールを使いこなしつつ、無駄な労力を省いて目的を達成できるようになる。機械学習アプリケーション作成に必要な機械学習のフレームワークやツールの経験が豊富な開発者ならば、こうした AMI を使い短時間でアプリケーションを導入することが可能になる。

課題と機会

機械学習を最初期から導入していなかった多くの企業は今、テクノロジーがビジネスに最大のベネフィットをもたらす分野を見極められずにいる。そればかりでなく、AI アプリケーションの構築と導入に必要なスキルセットについても確信が得られていない。要するに、各企業とも AI アプリケーション活用に適切なツールやインフラ、環境を模索している段階である。

AI アプリケーション開発に貢献できるツールや技術の特性について、各企業は何らかの指針を必要としている。また、自社に最も有効なアプリケーションがいつ、なぜ、どのように効果を発揮するのかを理解しなければならない。加えて、深層学習を採用した今後のプロジェクトへの投資効果（ROI）を評価するため、アプリケーションの効果を数値化する必要もある。

最後になるが、AI に関する Platform as a Service（PaaS）市場にはすでに多数の参入者があり、日々競争が激しくなっている。AI アプリケーション市場の主な関心事は優れた（かつ簡素な）ツール、市場投入の早期化、効率の必要性和需要である。この分野では実績があるものから誕生したばかりのものまで、さまざまなベンダーがサービスやソリューションを企画、提供しており、能力的にも幅が広い。そうした状況の中、AWS は指導的役割を維持できるか問われているため、急ピッチでエンジニアリングとイノベーションを続けていく必要に迫られている。AWS ソリューションの根底となる機械学習や深層学習の製品化自体は、市場では既知の手法である。ただし、AWS Deep Learning AMI や SageMaker などのマネージドサービスは、深層学習用の多数のツールとフレームワーク、テクノロジーを 1 つの統合プラットフォームにまとめ、企業や開発者の生産性を大幅に高めている点が新しいと言える。この新興市場に対し、AWS は今後も同じレベルのイノベーションと専門知識を投入し続けていかなければならない。

結論

機械学習や深層学習の導入を成功させる上で不可欠な要素は、人とプロセス、そしてテクノロジーである。これまで生まれてきた新しい技術的ソリューションは、最先端技術に生きがいを感じる先鋭的、意欲的な開発者の存在を必要としていた。それに対し、クラウドベンダーは AI や深層学習の開発と利用を開放し、企業での利用や実装を広める道を探り当てようとしている。深層学

習に基づいて短時間で開発したモデルや製品を成功させられることが鍵となる。これに役立つ要素を以下に挙げる。

- **クイックスタート用のパッケージと開発ツール**：一部のベンダーは開発者がスタートと同時に高い生産性を発揮できるように、テンプレートやサンプルデータ、サンプルコードを用意している。これに対し Amazon SageMaker のようなマネージドサービスを利用すれば、データサイエンティストや開発者（さらには開発者以外の要員）の生産性はテンプレートやサンプルコードにいつそう向上する。
- **データ面のサポート**：一部のベンダーは、サードパーティからのデータキュレーション（テーマを絞ったデータ収集）を検討、あるいは実行することで、ユーザーが求める最先端の予測型および処方型機械学習モデルの構築を目指す開発者を支援する。自前のデータを持つユーザーに対しては、多数のクラウドベンダーからデータキュレーションサービスが提供されており、データセットの入念な検討に貢献している。AWS もまたは画像、Web クロール、報道、Web 動画など汎用的な機械学習データの他にも各種テーマ（地理空間、環境、遺伝子工学、生命科学、法規制、統計データ）に特化した公開データを用意している。
- **教育**：ツールの最適用法に関する研修やトレーニングの機会を設けることで、開発者は何もかも自前で用意しなくてもすぐ作業に着手できるようになる。Amazon SageMaker などの機械学習や深層学習によって、簡単な教育だけで従来とは異なる層の開発者やデータサイエンティストの生産性を高め自分の力で深層学習モデルを構築、導入できるようになった。AWS の場合、アマゾン自社の開発者やデータサイエンティストのトレーニングと同じ機械学習トレーニングが無料で提供されている。開発者には AWS DeepRacer を使った実習コースも用意されている。これは 18 分の 1 スケールの完全自律型のレーシングカーで、自動走行を通じて開発者が強化学習について学べるよう設計されている。
- **コンサルティング/アドバイザリーサービス**：これは開発者の生産性を高めるほか、使用データに関するさまざまな課題の解決に貢献するサービスである。例として、AWS Partner Network に対応した新たな AI および機械学習のコンピテンシーが挙げられる。アマゾンはこのプログラムを取り入れた機械学習や深層学習のパートナーを認定している。加えて、AWS はアマゾン機械学習ソリューションラボ（Amazon Machine Learning Solutions Lab）を立ち上げ、AI アプリケーション開発の迅速化と簡素化を目指すユーザーを支援している。ソリューションラボではユーザー側と AWS 機械学習専門家が共同でデータの準備、モデルの構築とトレーニング、モデルの本番移行に取り組んでいる。

優秀な AI アプリケーションには、先進技術と堅実な設計上の判断が両方必要となる。各企業とも、今後構築する AI ソリューションが深層学習の採用によってビジネス上の目標達成と課題克服に役立つことを確認しておくべきである。社内の有識者や本来のステークホルダー、コンサルティングパートナーを巻き込んで、期待されているビジネス目標を達成できるよう適切なユースケースを策定しなければならない。組織内のデザイン検討過程に過去のプロジェクト経験を織り込むだけでなく、同業他社で開発された既存ユースケースも可能な限り取り込んで最終目標達成に最適なユースケースを確立すべきである。この過程では、正しいユースケースを打ち立てたとの確信が得られるまで、イノベーションとプロトタイピングを繰り返すことが求められる。

多種多様なツールやライブラリーが提供されている中、目的とするユースケースや業務に最適のサービスやライブラリーを探し出して AI アプリケーション開発を成功に導くことは容易ではない。Amazon SageMaker や AWS Deep Learning AMI などのソリューションは、開発者が生産性を高め、深層学習モデルだけでなくデータのキュレーション、統合、管理に関するサービスを導入することで、従来のコーディング手法では限界のあった課題を克服し自社のビジネス要求に応える上で有益である。各企業はこうしたツールやサービスを検討した上で深層学習モデルを使った AI 対応アプリケーションの開発、導入を進めるべきである。

IDC 社 概要

International Data Corporation (IDC) は、IT および通信分野に関する調査・分析、アドバイザリーサービス、イベントを提供するグローバル企業です。50 年にわたり、IDC は、世界中の企業経営者、IT 専門家、機関投資家に、テクノロジー導入や経営戦略策定などの意思決定を行う上で不可欠な、客観的な情報やコンサルティングを提供してきました。

現在、110 か国以上を対象として、1,100 人を超えるアナリストが、世界規模、地域別、国別での市場動向の調査・分析および市場予測を行っています。

IDC は世界をリードするテクノロジーメディア（出版）、調査会社、イベントを擁する IDG（インターナショナル・データ・グループ）の系列会社です。

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com

Copyright Notice

External Publication of IDC Information and Data – Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2019 IDC. Reproduction without written permission is completely forbidden.

