



最新 AI で飛躍的成長 を実現する

プラットフォーム時代に
AI 投資から画期的な効果を得るためのガイド



要約

どの業界においてもディストラクター（破壊的変革者）たちは、カスタマー・エクスペリエンスや市場投入速度、イノベーションに新たな標準を打ち立ててきました。人工知能（AI）活用は、トップ企業が画期的結果を実現して市場を再構築し、業界で際立つ存在となるという、1つの転換点に達しています。AIの中心にあるのは、自動化、予測、最適化という戦略的なイネーブラーです。日常業務を自動化し、成果を予測し、リソースを最適化する組織能力は、組織が成長するために不可欠です。実際に、高い成長を遂げる企業は、優れたカスタマー・エクスペリエンスの創出、製品とサービスの迅速な提供、運用の合理化、エコシステムへの投資といったビジネス上の必須要件だけでなく、コンプライアンスとリスク管理の要件も十分に満たしています。

本書では、以下の点を探ります。

- 高成長のトップ企業による AI 投資の特性
- データと AI のプラットフォームが必要な理由
- データと AI プラットフォームが目指すべきもの: 自動化、予測、最適化
- 信頼性と透明性を備えた AI を構築しスケーリングすることの利点

本書を読むことで、業界のトップ企業が AI をどのように活用しているのか、プラットフォーム・アプローチの重要性、データと AI のプラットフォームが生み出すメリットについて洞察を得られます。また、本ガイドでは、どのようなアクションを実行できるのかを中心に説明し、ビジネス成功に役立つ戦略を探ります。

「データのガバナンスと AI モデルの管理方法、それらをどのように検証して使用するのかは、現在、役員会レベルの課題となっています。もちろん、データの倫理的な使用についても同様です」¹

Banking
Netherland社、COO

AI のオポチュニティーと課題

企業は AI を使用して、ビジネス成果の予測、運用合理化、効率改善、サイバー脅威と不正行為からの防御、新しい市場オポチュニティーの発見を行っています。これらの予測は、トップ企業が競合他社や市場変動の一步先に行くのに役立ちます。また、AI の実装では、企業の役員はプロセス、テクノロジー、組織に重要な変更を加えながら、株主の期待に応えるというプレッシャーにさらされます。

さらに、AI モデルに関連するデータとガバナンスについて、役員会レベルの精査も行われます。[IBM Institute of Business Value は C-suite Study をリリースしました。](#)¹ タイトルは「信頼の優位性を構築 – あらゆる場所でのデータと AI 時代のリーダーシップ」で、ブランド名や機関に対する顧客からの信頼が急速に失われかけていることを示唆しています。お客様は製品とサービスに関連するデータの透明性を求めており、自らの個人データが安全に保管され、適正に使用されることの保証を望んでいます。



AI が勝利を導く高成長トップ企業

Forrester Research 社によると、成長率の高いトップ企業は AI に大きく投資しています。Forrester 社の調査に回答した企業の 50% 以上が、AI への投資の 5 倍を超えるリターンを得ることを期待しています。² これを大局的に見ると、1000 万ドルを投資した成長率の高いトップ企業は 6000 万ドルの ROI を見込んでいるということです。また、トップ企業は成長率の低い企業と比べてデータと分析の予算に 2 倍、AI と機械学習 (ML) のプラットフォームに 2.5 倍以上を投資しています。

Forrester Research 社はまた、本格的なスキルを持つデータサイエンティスト (予測の構築、ML、ディープ・ラーニング、自然言語処理 (NLP)、コンピューター・ビジョン、他のタイプのモデルに精通した専門家に投資した企業は、そのような投資をしていない企業よりも成長が早いことを発見しました。

Forrester 社によると、特定の AI アプリケーション (HR、セールス・リード・スコアリング、不正検出を含む経費管理) については、組織はパッケージ化されたソリューションの購入を好んでいます。

- 46% は、上記のようなアプリケーションの場合、AI が組み込まれパッケージ化されたソリューションを購入
- 20% は、社内で AI を開発

実際には、自社のビジネス・ニーズに基づいて、購入、構築、またはその両方を容易に行えるデータと AI のプラットフォームが必要です。

データと AI のプラットフォームを評価するときに考慮すべき要件

データサイエンスは、企業が意味のあるパターンを認識し、成果を予測し、意思決定を簡素化するのに役立つ分野であり、AI を活用する上でのキー・アクセラレーターです。データから得られた新たな洞察やパターン、その他の価値ある発見を活用することで、企業は次に何が起きるのかを予測して意思決定を簡素化できるようになります。適切なオファーとアプローチを組み込んでアクションを最適化すれば、選択したシナリオに基づいた可能な限り最善の成果達成を追求できます。このポテンシャルこそが、アイデアを実行し、データサイエンスを積極的に活用するためのプラットフォームが必要となる理由です。

では、プラットフォームとは何でしょうか。プラットフォームは、「参加者の間で価値創造の相互作用を促進するインフラストラクチャーです。プラットフォームは、このような相互作用にオープンな環境を提供し、そのためのガバナンス条件を設定します。」³ AI を用いてビジネスの最新化を成功させるには、適切なプラットフォームを選択することが戦略的に不可欠です。

人材、データ、信頼に取り組むための基盤を構築

企業の目標は、予測と最適化のプロセスを反復的な変革およびインテリジェントなワークフローに変換することです。この目標を達成して組織内で AI を発展させるためには、従業員に対する開発の自動化に役立つデータと AI のプラットフォームに関する簡単な研修が必要になります。さらに、専用ツールを追加するのではなく、新規プラットフォームを使用して既存のテクノロジー投資を活用する方法を見つける必要もあります。

人材: AI の成功にはスキルとスタッフが不可欠です。PwC の「2019 年版 AI 予測」に示されているように、AI に対応した従業員が必要です。これには、人材の再教育とスキルアップのための継続的な学習の取り組みが求められます。さらに、職務内容が変更された場合は、従業員の訓練の進化に合わせ組織の構造を見直す必要もあります。⁴

データ: データはモデル・パフォーマンスの生命線です。AI モデルはデータの上に成り立っており、モデルがどのようなパフォーマンスを発揮するのかは適切な実動データによって決まります。この仕組みは、プラットフォームで継続的なデータ・フローへのアクセスを可能にする必要があることを意味します。また、トレーニングされたデータとは異なる実動データにモデルが遭遇した場合は、精度のずれは避けられないので、これを検出し緩和する必要があります。そして、AI のライフサイクル全体で使用されるモデルとデータに関し、監査可能な記録を提供する必要があります。

信頼性: チームは、AI モデル内でのバイアスをどのように検出して緩和するのかを示し、個々の成果を説明できなければなりません。また、プラットフォームでは、ビジネス KPI に基づき成果を追跡できる必要があります。さらに、AI 関連の取り組みを拡大して持続できるように、プラットフォームには、確実な信頼性と説明可能性が組み込まれていなければなりません。

AI による予測と最適化をアプリケーションに迅速に統合

データと AI のプラットフォームもまた、AI モデルから最新アプリケーションへの洞察の統合をサポートすべきです。アプリケーション開発には、ほとんどの企業が既に大きく投資しています。データと AI の柔軟でオープンなプラットフォームは、アプリケーション開発チームとビジネス・チームがモデル運用 (ModelOps) を構築するときの基盤として機能します。ModelOps は DevOps とシームレスに連携し、AI を用いた最新アプリケーションを成功に導くための支援をします。

AI ライフサイクル管理を自動化

従来、データサイエンスと AI への投資は、予測分析と ML を活用して、ビジネス上の疑問点に回答したり、プロセスの小規模な部分を自動化することに重点を置いてきました。しかし、現在、トップ企業の多くは、AI の活用範囲を広げることに注力しています。この動向は、企業の隅々までモデルとツールの管理を運用可能にして自動化できるようにプラットフォームを設計することを意味します。

自動化によりチームは、企業を差別化するための価値のあるアクティビティに再び集中できるようになります。以下のようなステップを自動化できるプラットフォームを探します。

- データの準備
- フィーチャー・エンジニアリング
- 機械学習アルゴリズムの選択
- ベストな ML モデルを選択するハイパーパラメーターの最適化

この一連のステップは、プロセスの各ステージにおける最も有望なステップを進む AI システムのガイドに沿って進める必要があります。これが AI を使用した AI の構築であり、例として、IBM Research を採用した機能の AutoAI があります。⁵

予測的な成果に基づいて意思決定を最適化

高成長のトップ企業の中には、複数のユース・ケースに取り組み、AI を使用して意思決定を向上させるスキルを持っている企業もあります。AI の洞察と AI のオーケストレーションは人材と組み合わせることができます。企業が最高の ROI を達成するには、データと AI プラットフォームで予測的な成果を使用し、それを用いてアクションを設定する必要があります。

チームの最適化の課題に対処するために、データと AI の最新プラットフォームではデータを選択して編集するワークフローを促進する必要があります。このプラットフォームでは、自然言語のインターフェースを使用して、チームのメンバーが最適化モデルを実行し、ガント・チャート、スケジュール、リソース計画、需給の割り振りなどを用いてレポートを作成して共有する必要があります。データと AI のプラットフォームの中で意思決定を最適化すると、予測的な成果への処方的分析の適用が簡素化されます。これが意思決定のインテリジェンスです。⁶

多様なスキルと才能を持つ従業員のスキルアップと再教育によって生産性を向上

高成長企業で成功しているデータ・サイエンス・チームには多様性があります。計算、データ管理、応用数学、ビジネス・ユースケース、意思決定サイエンスなど、複数の AI 分野で専門家になれる人は一人もいないでしょう。AI には、さまざまなスキルを持つ人々によるチームと、コラボレーションするという視点が重要です。

このために、視覚的なデータ・サイエンス・ツールとオープン・ソース・ツール、そして Python や R などのプログラミング言語を使用して、モデル構築のための視覚的なアプローチとプログラマチックなアプローチの両方をサポートできるプラットフォームが必要になるのです。また、このプラットフォームでは、ビジュアル・データ・サイエンティストが迅速にテキストと構造化データを組み合わせる隠れたパターンを明らかにし、このパターンを予測と最適化に活用できるよう支援する必要があります。

ビッグデータへの既存の投資を組み合わせる活用

より安全かつ高速に大きな洞察を得るためには、プラットフォームで分析とビッグデータを近づけることも重要です。例えば、データと AI の最新プラットフォームには、大量のデータを移動せずに、Hadoop クラスターからデータにアクセスしたり、Hadoop クラスターに実行をプッシュしたりするセキュアな方法が用意されています。企業と規制当局は、潜在的なバイアスを緩和して成果を説明することを求めています。したがってこれらの機能をデータと AI の最新プラットフォームに組み込まなければなりません。

AI モデルの信頼性と透明性を向上

データと AI モデルに信頼性を浸透させることが最重要です。これを成功させるには、AI のライフサイクル全体で、AI の成果を追跡して測定する必要があります。また、モデルが作成された場所や実行されているクラウドにかかわらず、モデルが公正かつ説明可能で規制に準拠しているという状態を維持するための機能も必要です。企業と規制当局は、潜在的なバイアスを緩和して成果を説明することを求めており、それゆえこれらの機能をデータと AI の最新プラットフォームに組み込まなければなりません。

クラス最高水準のオープン・ソース・ツールで構成されるエコシステムをあらゆるクラウドに提供

チームはどこにでも存在し、データもまた同様です。世界中で発生しているイノベーションを利用するには、データがどこにあってもその場所にモデルを用意する必要があります。データと AI プラットフォームはオープンでなければならず、複数のクラウドで稼働するモデルとデータをサポートするとともに、活力あるエコシステムからメリットを得る必要があります。また、プラットフォームでは、データ移動の必要性にともなうコストとリスクを緩和できるようにする必要もあり、これによって規制や法的懸念が生じる場合もあります。

さらに、プラットフォームには、業界のアクセラレーターを使用して AI プロジェクトを一気に開始する機能や、事前定義されたビジネス用語とデータ・サイエンスの成果物を用いて再構築したアプリケーションも必要です。⁷

柔軟なライセンス・モデルでの構築と購入の選択の組み合わせ

テクノロジーに関して、構築するか購入するかは判断は、しばしば複雑な課題です。高成長企業のほとんどは、データ・サイエンスと AI という観点から、購入すべき、または構築する時期を決定しています。ライフサイクルが進むにつれニーズは変化します。このため、パッケージ化された AI アプリケーションに対応できるだけでなく、オープン・ソース・テクノロジーと専用テクノロジーを融合したエンタープライズ・クラスのツールを用いた構築もサポートできる、データと AI の柔軟なプラットフォームを探する必要があります。

アジャイルな AI プラクティスを統合ツールと連携

データ・サイエンスと AI の実力を引き出すために、組織はアジャイルな AI を実装する必要があります。数百もの AI 実装の成功事例から学んだベスト・プラクティスを活用できます。⁸

このような基本的な技術原則を検討すると、AI に関してビジネス・リーダーらを指導するときに役立ちます。高成長企業がどのように AI を使用しているかを説明すれば [\(外部ページヘルプ、英語\)](#) コラボレーションが、個人レベルおよび組織レベルの生産性の大幅な向上をどのようにもたらすかを強調できます。AI のセンター・オブ・エクセレンス (CoE) にシニア・リーダーシップ・サポートを維持し提供することで、AI 投資から高いリターンを確保することに役立ちます。

Watson Studio Premium for IBM Cloud Pak for Data

Watson Studio Premium for IBM Cloud Pak for Data は、AI 投資による価値実現までの時間短縮に寄与します。Watson Studio Premium は、IBM Decision Optimization、SPSS Modeler、Hadoop Execution Engine で構成されています。

この強力な組み合わせは、以下の点で企業に役立ちます。

- データと AI のプラットフォームの一部として、予測的な成果に基づいて意思決定を簡素化する
- 場所を問わず、データと分析の専門家を支援する
- ビッグデータへの既存の投資を活用する

このプラットフォームにより、組織は単一の統合環境でビジネス成果を独自の方法で予測し最適化できるようになります。

IBM Cloud Pak for Data は Red Hat OpenShift 上に構築され、わずか数時間で導入できるとともに、IBM とサード・パーティー製の増え続けるマイクロサービス群を用いて容易に拡張できます。IBM Cloud Pak for Data はあらゆるクラウドで稼働するので、組織は分析とアプリケーションを容易に統合し変革を速めることができます。企業はこの最新の AI プラットフォームで AI と ML のモデルを構築して展開し、データ・サイエンス・プロジェクトを加速し、自動化されたエンドツーエンドのライフサイクルを実現することができます。IBM Cloud Pak for Data は、Amazon Web Services (AWS)、Azure、Google Cloud、IBM Cloud、プライベート・クラウドなどのマルチクラウド環境をフルにサポートします。AutoAI と Watson OpenScale を、IBM Cloud Pak for Data の基盤の一部として使用することもできます。



業界共通の AI ユース・ケース

コール・センターの効率を向上 お客様の電話、Eメール、SMS、チャット・メッセージの予測不能な急増に対応できるチャットボットを使用します。

クロスセルとアップセルを増加 パーソナライズされたリアルタイムの推奨と提案を行います。

ロイヤリティの向上 顧客の解約や乗り換えを予測し、引き留めるための手順を推奨します。

オフリングを最適化 お客様の声に耳を傾け、将来のニーズを予測します。

マーケティングの強化 ターゲットを絞って、パーソナライズされたキャンペーンを実施します。

在庫コストを最小化 正確な予測で、リソース管理を向上させます。

生産性の向上 適切な従業員を適切なタイミングで適切な職務に割り振り、労働力を正確に予測します。

保守コストを削減 発生前に障害を予測します。

リスクを緩和 お客様の正確な信用評価を行います。

不正行為の検知 疑わしい行動パターンを特定します。

新規ビジネス・モデルを活用 埋もれた需要を掘り起こし、予測を最新アプリケーションに統合します。

信頼性と透明性を備えた AI を構築しスケーリングすることの利点

AI の成功に向けて、組織を率いて先導しています。AI を収益化するには、データと AI の適切なプラットフォームに投資する必要があります。適切なプラットフォームとは、データ、人材、信頼性に関する課題に変革的な方法で対処できるプラットフォームです。このプラットフォームでは、予測、自動化、最適化を AI 実装の一部として統合する必要があります。信頼性と透明性を備えた AI を構築してスケーリングすることにより、以下の領域で AI のパワーを活用できます。

- AI ライフサイクル管理の自動化
- 予測的な成果を使用した、意思決定インテリジェンスの向上
- モデル・パフォーマンスのモニタリング
- 統合マルチクラウドによるデータと AI プラットフォームとの協調的データ・サイエンス促進

これら利点は大きなものです。組織は、収益、採算性、優れた運用、変革の面で、同業他社を上回ることができます。信頼性を向上させ、差別化された手法を用いて顧客とビジネス・パートナーを引きつけることのできる方法で、データとモデルを処理することができます。また、データと AI モデルを信頼し、その信頼できるモデルを使用して顧客とパートナーの体験をバリュー・チェーンに沿って高め、成長を加速させることができます。

詳細情報

IBM Cloud Pak for Data によって、お客様のビジネスをどのように変革できるのかをご覧ください。

[仕組みを見る](#) →

Web セミナー: AI で勝利するためのプレイブック

AI 投資によるリターンを高める方法と、トップ企業のどこが違うのかを 3 部構成の Web セミナーで説明します。

[Web セミナーに登録する \(外部ページヘリンク、英語\)](#) →

エンタープライズ AI の ROI

IBM の統括責任者兼最高データ責任者である Seth Dobrin が、探索と調査のためのエンタープライズ AI から、ビジネスに不可欠なエンタープライズ AI へと移行することで、真の ROI を取得する方法を説明します。

[動画を見る\(2:08\) \(外部ページヘリンク、英語\)](#) →

ビジネス価値を評価

改善されたアクションを提供する信頼された AI ソリューションには、大きなビジネス価値を生み出す可能性があります。AI によるビジネス価値の評価は、どれだけの価値を生成できるのかを示します。

[アセスメントを始める \(英語\)](#) →

製品ウォークスルー

AI の価値向上の一環として、データが確実に高品質となるので、自動化された正確な洞察と意思決定を得られます。IBM Cloud Pak for Data なら、それに貢献することができます。製品の詳細をご覧ください。[製品ウォークスルーを視聴する \(外部ページヘリンク、英語\)](#) →

付録: 業種別 AI コース・ケース

高成長のトップ企業は、ビジネス上の特定の問題を徹底的に解決するために、自社の AI モデルの構築にどのように投資しているのでしょうか。業種別に、以下の例を検討します。

銀行

- 取引に関する市場と取引先のリスク評価
- ローン申請の信用リスク評価
- 不正取引のリアルタイム検知
- 信用リスクの迅速な評価によってリスクを制御しながら、ローン審査承認と顧客サービスの加速

保険業界

- 不正請求の検知
- 申込者ごとに関連リスクを評価することで、見積りと保険料の最適化
- 災害の可能性がある気象変化を予測して、自動車保険の請求を削減

公益事業

- 現物資産の膨大なネットワークの管理
- 生産と需要のパターン予測
- 発生前の障害予測
- 需要供給計画
- 需要と供給の最適化プロセスの合理化

政府・官公庁・自治体

- 利益に関する不正検知
- 公共サービスの利用パターン予測
- 廃棄物管理と交通量の最適化

製造業

- 製造品質をモデル化して欠陥を検知することで、製造ラインの円滑な稼働状態の維持
- ウェアハウス管理と物流の最適化
- ML モデルを使用して、自律走行車両のセンサー開発
- サプライ・チェーンの運用最適化

小売業

- カスタマー・ロイヤルティ・プログラムの管理
- 顧客プロファイルと高度な傾向モデルに基づいてターゲットを絞った推奨を作成することで、クロスセルとアップセルを強化
- 正確な需要予測の実現

食品業

- 食品の安全性に関するデータ収集と分析の自動化
- 健康被害の可能性を予測して警告し、迅速な介入の実施
- 競合他社と安全にコラボレーションできるよう、機密情報を保護

医療

- ECG や他の医療デバイスからのデータの流れを監視
- 患者の容態が変化する可能性を予測
- 医療調査を実施
- 患者データの流れのリアルタイム分析

メディアとエンターテインメント

- TV視聴者に関する、より速くより深い洞察の提供
- 詳細で複雑な視聴者データに関する洞察力の強化
- 柔軟なリソースを使用して、需要の変化に応じて迅速かつ容易にスケールングを実行
- テクノロジーや運用よりも、ビジネス活性化に集中

教育

- 生徒の達成や定着を予測
- 目標達成のために追加サポートが必要な生徒の特定
- ドナー・リレーションシップ強化
- 生徒の動向を追跡で、欠席を減らせるよう支援

コンピューター・サービス業

- 企業の作業プラクティスに関する即時の洞察の提供
- ロボティック・プロセス・オートメーション (RPA) を用いて卓越した効率を実現
- 従業員がより戦略的な活動や業務に集中できるよう支援
- 顧客行動予測とチャーン・モデルに基づいて優れたサポート体験を作成し、顧客満足度を向上

用語集

アルゴリズム は、特定の問題を解決するためにデータに適用できる一連の運用を定義するルールのセットです。

人工知能 (AI) は、データを解釈してデータから学習するコンピューター・システムの能力です。この用語が最も一般的に使用されるのは、機械学習モデルまたはディープ・ラーニング・モデルを使用して構築されたシステムを表すときです。AI の技法は、これまで解決が困難だと思われていた幅広い問題をコンピューターが解決できるようにするために使用できます。

バイアス は、モデルを設計、トレーニング、テストするときの一般的な問題であり、不正確な予測を導く可能性があります。企業は AI の採用をより広範囲に検討するようになり、実行時のモデルのモニタリングと監査によるバイアス軽減は、ますます重要なトピックとなっています。

分類モデル の目的は、既にカテゴリに分類されているデータ・ポイントのセットと比較することで、データ・ポイントをカテゴリに分類することです。結果はスコアではなく、限定されたオプション・リストの 1 つを示す離散値になります。例えば、顧客が購入する可能性があるかどうか、信用リスクが懸念されるかどうかといった質問に対し、分類モデルは「はい」あるいは「いいえ」のいずれかの回答をします。

コンテンツ分析 は、さまざまな形式のドキュメントに含まれる非構造化データ (テキスト、イメージ、オーディオ・ファイル、動画ファイルなど) を分析することです。これまで、確認と分類に何百時間または何千時間も要していた大量コンテンツが含まれるリポジトリ分析が、機械学習の技法によって大きく加速されました。

データ・サイエンス は、統計、データ分析、機械学習の側面を統合する幅広い分野であり、データを活用してビジネス課題を解決します。新しいパターンの認識、予測の構築、意思決定の最適化に役立ちます。

意思決定の最適化 は、数学と人工知能の高度な技法を使用して、数百万もの決定要因、ビジネス上の制約、トレードオフがからむ意思決定の問題を解決します。

ディープ・ラーニング は、多数の隠れたレイヤーを持つニューラル・ネットワークを使用する機械学習のひとつの分野です。ディープ・ラーニングの最先端の分野 (コンピューター・ビジョン、機械翻訳、音声認識など) には、非常に高度なネットワークが使用されます。

デプロイメント は、モデルをビジネス・アプリケーションに統合し、現実のデータに対してそのモデルを実行するプロセスです。テスト、ステージング、実動の各環境でモデルを作成して移動するには、データ・サイエンス、アプリケーション開発者、IT 運用の各チーム間でのコラボレーションが必要です。

予測モデリング では、従来型の統計技法または機械学習アルゴリズムを使用して、データ・セットに対するトレーニングとテストを実行し、モデルを作成および改良します。開発プロセスは、非常に多くの反復で成り立っています。必要なレベルの正確さを達成するまでに、数十のモデル、ときには数百のモデルをトレーニングしなければならないことがあります。

説明可能性 は、意思決定ごとにコンテキストを提供して、AI モデルの透明性を確保し、監査可能にします。これは、予測モデルを使用して推奨事項を作成し、ビジネス上の意思決定を支援するシステムにとって重要な属性です。複雑でミステリアスに見える予測モデルでは、企業のステークホルダー、規制当局、顧客の信頼を勝ち取ることはできません。

データ探索 は、モデル構築プロセスの重要な部分です。このアクティビティの目的として、特定のデータ・セットの興味深い特性を見つける、隠れた関係性を明らかにする、予測モデリングがビジネス価値をもたらすであろうユース・ケースに焦点を当てるなどが挙げられます。

地理空間分析 は、緯度と経度、郵便番号と住所などの地理データ分析です。この分析は、データ・サイエンスに関する多数の種類の実用的な問題を解決するために非常に役立ちます。最新のデータ・サイエンス・プラットフォームは、地理空間情報の検出、解析、計算を容易にするだけでなく、マッピング・ツールと簡単に統合でき、結果を視覚化することもできます。

推論 (人工知能における推論) は、論理的なルールを知識ベースに適用し、不確実性の中で結論を引き出します。ユーザーは、実行時のパフォーマンスに関連する、簡素化、圧縮、最適化について、推論を用いた予測を得られます。

線形回帰 は、1 つの独立した変数を使用して、値やスコアを説明または予測する統計プロセスです。例として、特定の週に販売された製品の SKU の数や、アカウントを閉じる顧客のリスクのパーセンテージなどがあります。

ロジスティック回帰 は、成果の予測に使用される統計プロセスです。このプロセスが線形回帰と異なるのは、1 つの独立した変数にとる値は、無限にあるのではなく限られた数しかないという点です。ロジスティック回帰は、応答がカテゴリ (例えば、1 番目、2 番目、3 番目など、数値による順番を持つカテゴリ) に分類される場合に使用されます。

機械学習 (ML) は、統計技法を使用して、明示的なプログラミングを必要とせずに、大量データ・セットから精巧な予測モデルやアルゴリズムを導き出します。一般に、トレーニングとテスト用にデータ・セットを 2 つのサブセットに分割することから、この反復プロセスを開始します。数十または数百のバリエーションを用いて、トレーニング・セットに対するモデルのトレーニングと、テスト・セットに対するモデルの性能テストを行うことで、予測精度を評価します。このプロセスを実行し、それぞれの反復で最高の性能を示したものを次世代バリエーションの土台とすることで、モデルは徐々に学習し、性能が向上します。

モデル管理 は、経時的にモデルの正確さを保つために不可欠です。モデルを定期的に再トレーニングして新規データを考慮に入れることが重要なので、モデルの開発、実装、導入、管理を継続的なサイクルにする必要があります。

ModelOps (モデル運用) は、AI モデルをアプリケーション開発に統合して導入し、DevOps とアプリケーション開発ライフサイクルを拡張する機能です。運用化するモデルは ML モデル、意思決定最適化モデル、データ・トランスフォーメーション・モデルのいずれでもかまいません。モデル運用は、従来型の ML モデルだけに限定されず、さまざまなモデルで構成できます。

自然言語処理 (NLP) は AI のひとつの分野であり、主に、構造化されていないテキスト・データをコンピューターが分析できるようにすることに焦点を当てます。一般的なユース・ケースとして、音声認識、自然言語理解、センチメント分析などがあります。

ニューラル・ネットワーク は、多数の機械学習アルゴリズム間の複雑な相互作用を可能にして最適なモデルの特定を支援するトレーニング・モデルのフレームワークを提供します。人間や他の動物の脳内にある相互接続されたニューロンの構造が、人工的なニューラル・ネットワークの構造を生み出すきっかけになりました。人工的なニューラルはレイヤーによって接続されます。データは、入力レイヤーから、1つ以上の隠れたレイヤーを経由して、出力レイヤーへと構造内を通過していきます。この通過中に、数学関数が精度を評価できる予測へとデータを変換します。

オープン・ソース・ソフトウェア は、統計モデリングと機械学習の多数の領域で、ますます有力なパラダイムになっています。R、Python、Scala などの言語、Apache Hadoop や Spark などのビッグデータ・アーキテクチャー、TensorFlow や Spark MLlib などの機械学習フレームワークはすべて、予測分析とデータ・サイエンスの世界で主役となっています。

予測分析 は履歴データを使用して特定の領域または問題をモデル化し、これまでに特定の成果を上げてきた主要要因を分離します。このプロセスを使用して構築されたモデルは、新しいデータから、考えられる将来の成果を予測します。予測分析には、従来の統計モデリングから機械学習アルゴリズムに至るまで、幅広い技法を含めることができます。

予測モデル は、入力を出力にマッピングするアルゴリズムです。この場合の入力は1つのデータ（データベースのレコード、テキスト・サンプル、イメージなど）を意味し、出力は予測となります。一般に、出力は連続変数（数値またはパーセントなど）か、離散カテゴリー（「はい」、「いいえ」など）のいずれかです。予測モデルには、回帰モデルと分類モデルの主に2つのタイプがあります。

データの準備 は、データ・サイエンス・プロセスの最初のステップの1つです。ほとんどのプロジェクトは、データ・セットを改良し、詳細な分析の負担に耐えられるよう十分高品質にすることから開始します。多くの場合、ソース・データをクリーンにして、モデリングや分析に適した形式に変換する必要があります。機械学習モデルを構築している場合は、教師あり学習で使用するために、手動でのデータのラベル付けが必要になることもあります。

回帰モデル は、複数の変数を持つデータ・セットがあり、その変数間の関連性を分析したいときに役立ちます。特に、回帰モデルは、他の変数が変更されたときに、特定の1つの変数にどのような変化を及ぼす可能性があるのかを明らかにできます。線形回帰は、値やスコアの予測に使用できます。例として、特定の週に販売予定の製品の最小在庫管理単位 (SKU) の数や、アカウントを閉じる顧客のリスクのパーセンテージなどがあります。

統計モデリング は数学のひとつの分野で、データ・セットに関する確率的な仮定に基づくモデルの作成に使用されます。企業は統計的モデルを使用してデータ・セットの重要な特性を分析し、データの分類または予測の生成に使用できる相関関係を特定しています。

教師あり学習 は、データが既に正しくラベル付けされたデータ・セットを使用して機械学習モデルをトレーニングする方法です。モデルは出力変数（通常はカテゴリーまたは値）を生成するため、ラベル付けされた入力と出力とを比較することで、モデルの精度を簡単に評価できます。線形回帰、ランダム・フォレスト、サポート・ベクトル・マシンは、いずれも教師あり学習アルゴリズムの一般的な例であり、予測モデルのほとんどはこれらの技法を使用して構築されます。

テスト (予測モデルのテスト) は、トレーニングと同様に、AI プロセスでのデータの精度を判別するために不可欠です。予測モデルは継続的にテストし、精度を向上させる必要があります。モデルが失敗に終わった場合、アナリストは根本原因を特定し、再トレーニングとテストを行ってモデルを改善する必要があります。

テキスト分析 は、言語ルール、自然言語処理、機械学習を使用して非構造化コンテンツを測定します。このプロセスは、人間の頭脳に似たアプローチでデータを検討しますが、より高速にそれを行います。テキスト分析を用いると、全データの約90%を占める非構造化コンテンツからより多くの洞察と発見を得られます。

トレーニング (予測モデルのトレーニング) は、機械学習、ディープ・ラーニング、および有用なデータを見極めるためのその他の AI プロセスにとって重要な要素です。精度の高い予測を生成するようにトレーニングされたモデルは、リアルタイム・データのスコアを付けるために使用できます。モデルは定期的に再トレーニングし、変化する動作パターンに合わせて調整する必要があります。

監視なし学習 は、ラベル付けしていないデータを用いて機械学習モデルをトレーニングする方法です。データ内の興味深いパターンや構造をモデル化して強調する場合によく使用されます。クラスタリングとアソシエーションの問題は、非構造化学習の一般的な領域です（例えば、顧客を区分するための興味深い新たな方法を見つける、顧客間の類似点を特定するなど）。

視覚化 は、データをグラフィカルに表現するプロセスで、図やグラフを使用することがよくあります。人間がデータを理解するには、データを視覚化できなければなりません。このプロセスは、企業のステークホルダーに結果のプレゼンテーションを行うとき、プロジェクトの初期段階で新規データ・セットを探索するときのどちらにも重要です。

日本アイ・ビー・エム株式会社
〒103-8510
東京都中央区日本橋箱崎町19-21

IBM のホーム・ページは以下をご覧ください。

ibm.com

IBM、IBM ロゴ、ibm.com、IBM Cloud、IBM Cloud Pak、IBM Research、SPSS、Watson、および Watson OpenScale は、世界の多くの国で登録された International Business Machines Corporation の商標です。他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。現時点での IBM の商標リストについては、<http://www.ibm.com/legal/copytrade.shtml> をご覧ください。

本書の情報は最初の発行日の時点で得られるものであり、予告なしに変更される場合があります。すべての製品が、IBM が営業を行っているすべての国において利用可能なものではありません。

IBM 以外の製品とプログラムの操作またはサービスの評価および検証は、お客様の責任で行っていただきます。

本書に掲載されている情報は特定物として現存するままの状態を提供され、第三者の権利の不侵害の保証、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任なしで提供されています。IBM 製品は、IBM 所定の契約書の条項に基づき保証されます。

お客様は自己の責任で関連法規を遵守しなければならないものとします。IBM は法律上の助言を提供することはいたしません。また、IBM のサービスまたは製品が、お客様がいかなる法規も遵守されていることの裏付けとなると表明するものでも、保証するものでもありません。

適切なセキュリティの実施について: IT システム・セキュリティには、企業内外からの不正アクセスの防止、検出、および対応によって、システムや情報を保護することが求められます。不正アクセスにより、情報の改ざん、破壊もしくは悪用を招く場合があります、またはシステムの損傷や、他のシステムへの攻撃を含む悪用につながる場合があります。完全に安全と見なすことができる IT システムまたは IT 製品は存在せず、また単一製品またはセキュリティ対策が、不正アクセスを防止する上で、完全に有効となることもありません。IBM のシステムおよび製品は、包括的なセキュリティの取り組みの一部となるように設計されており、これらには必ず追加の運用手順が伴います。また、最高の効果を得るために、他のシステム、製品、またはサービスを必要とする場合があります。IBM は、何者かの悪意のある行為または違法行為によって、システム、製品、またはサービスのいずれも影響を受けないこと、またはお客様の企業がそれらの行為によって影響を受けないことを保証するものではありません。

© Copyright IBM Corporation 2020

- 1 IBM Institute for Business Value: IBM Global C-suite Study、第 20 エディション、「信頼による卓越 AI/Data 包摂時代のリーダーシップ」2019 年 11 月。
- 2 Forrester Research、IBM に代わって Forrester Consulting が実施した委託調査 (近日公開予定)、2019 年 8 月。
- 3 Sangeet Paul Choudary、「プラットフォーム・レボリューション PLATFORM REVOLUTION 未知の巨大なライバルとの競争に勝つために」2016 年。
- 4 PwC、「2019年AI予測AIを活用するための6つの優先課題」pwc.com、2019 年。
- 5 Julianna Delua、「AutoAIがAlconics Intelligent Automation Awardを受賞 – 第一人者が語る開発のねらい」ibmbigdatahub.com、2019 年 9 月。
- 6 Sajan Kuttappa、「The rise of decision intelligence: AI that optimizes decision-making (英語)」ibmbigdatahub.com、2019 年 10 月。
- 7 Paul Kilroy、「Industry Accelerators for Cloud Pak for Data – overview (英語)」ibm.com、2019 年 7 月。
- 8 Carlo Appugliese, Paco Nathan, William S. Roberts、「Agile AI: A Practical Guide to Building AI Applications and Teams (英語)」O'Reilly eBooks、2019 年。