



データサイエンティストとは何か。 本当にAIに取って代わられるのか？

～人気の背景と、AI時代の役割～

現在は瞬く間に膨大なデータが蓄積される時代です。

ビジネスや行政、医療、学術など多岐に亘る分野で、ネット上の購買履歴やSNSの情報、GPSの位置情報など、人や物のあらゆるビッグデータをいかに有効に活用できるかが、個人情報など倫理的な議論も含めて課題とされています。

とりわけ、ビジネスの分野では、ビッグデータの利用の巧拙の差が競争力の差につながるため、データ分析を専門とするデータサイエンティストの需要が高まっています。その一方で、AIの急速な進歩が、データサイエンティストの存在を脅かすであろうとも言われています。

そこで今回は、データサイエンティストについて、注目されるようになった背景とAI時代の役割についてみていきます。

データサイエンティストとは

データサイエンティストとは、ビジネスなどにおける課題解決や意思決定を、データの分析によってサポートする職種です。

そのためデータサイエンティストには、SASなどの環境や、R、Python、Goなどの言語を用いるITスキルはもちろん、統計解析や機械学習モデルの開発を行うスキルに加え、ビジネスの仕組みや市場動向に対する幅広い知識が求められます。

また、データサイエンティストはビッグデータを分析するための環境作りも行います。新商品やサービスを開発する場合には、どのような機能が、どのようなタイミングで必要とされるのか、消費者に関するさまざまなデータから分析します。

この分析結果が商品やサービスの競争力に大きな影響を与えるため、企業は優秀なデータサイエンティストの採用に注力し始めています。

データサイエンティストが登場した背景

データサイエンスとは、大量のデータから法則や関連性など意味のある情報を導き出す行為やその手法に関する研究を示します。

データサイエンスは2000年代に入ってから使われるようになった用語とも言われますが、最初に使われたのはデンマークのコンピュータ科学者であるピーター・ナウア (Peter Naur: 1928年10月25日～2016年1月3日) の著書『Concise Survey of Computer Methods』とも言われています。

このデータサイエンスの研究者や技術者をデータサイエンティストと呼びますが、注目されるようになったのは、2012年にHarvard Business Reviewがデータサイエンティストを「21世紀で最もセクシーな職業」([Data Scientist: The Sexiest Job of the 21st Century](#))と報じたことがきっかけになったと言われています。

このころ既に、企業には大量のデータが蓄積されていましたが、それらのデータは十分に活かされているとは言えない状況だったのです。

しかし、そのデータを活用することの重要性は認知され始めていましたので、データサイエンティストに対する需要が高まってきました。

データサイエンティストとは何か。本当にAIに取って代わられるのか？

ビッグデータとは

データサイエンティストは、さまざまな形式で蓄積されたビッグデータを分析します。

ビッグデータには、具体的なデータ量の定義はありませんが、その特徴として、データの発生頻度と多様性が高いことが上げられます。たとえばSNSで発信されている情報やスマートフォンから得られる位置情報、あるいは監視カメラの映像などがあります。今後はIoTの普及により、あらゆる製品のセンサーが捉えた情報も蓄積されていきます。

つまり、ビッグデータとは呼びますが、その特徴は単に量的なものにとどまらないということです。

データサイエンティストに求められていること

それでは、このようなビッグデータを扱うデータサイエンティストはどのような業務を行っているのでしょうか。

データサイエンティストは、PythonなどのAI・機械学習プログラミングに適した言語でデータマイニングを行うためのプログラムを開発し、実際にデータマイニングを行います。

データマイニングとは、統計学やパターン認識などのデータ解析手法を用いて、大量のデータから法則性や関連性といった意味のある情報を抽出する技術です。

また、AIによる機械学習やディープラーニングの活用方法を考案したり、データマイニングの結果を他部門で活用するための解析方法を提案したりすることもデータサイエンティストの役目です。

さらに、他部門のニーズを汲み取るためのコミュニケーションを取り持つ役割や、データ解析の方法や解析結果を経営部門が活用するためのサポートも行います。

データサイエンティストが求められている場

それでは、データサイエンティストを必要としている現場はどこでしょうか。

それは、システム内外に問わずビッグデータを保有・管理していながらも、それが経営課題の解決や経営戦略の立案に活用できていない企業です。

そのような企業はさまざまな業界に存在していますが、分かりやすいのはBtoC企業であると言えます。BtoB企業でも、消費者の動向を把握する必要がある広告関係やマーケティング関係が含まれます。

このような現場では、一般消費者の行動を、年齢別、性別、地域別、職業別、学歴別などあらゆる属性で分析することで、マーケティングデータとして分析することができます。

この分析結果を基に、製品やサービスの開発を行ったり、キャンペーンを行ったり、あるいはさまざまな告知方法を検討することができます。

データサイエンティストとは何か。本当にAIに取って代わられるのか？

データサイエンティストにはどのようなスキルが必要か

このように活躍の場が多いデータサイエンティストになるためには、幅広い知識とスキルが必要とされますが、ここでは主なスキルについて取り上げてみます。

●IT言語スキル

データサイエンティストはビッグデータを扱いますので、高いITスキルが求められます。

たとえばビッグデータの高速処理を行うためのApache Hadoopやデータベースの操作を行うための各種SQL言語、AI・機械学習に適したプログラミング言語のPythonやR、Juliaなどの知識が求められます。

●数学・統計解析スキル

プログラミングの知識を持っていても、データを分析するためには数学や統計学の知識が必要になります。

●ビジネススキル

データサイエンティストにはビジネスの課題を解決したり、経営戦略上のサポートを行うことが期待されるため、ビジネス全般や、所属する企業のビジネスモデルに対する知識と理解が必要になります。また、マーケットに対する知識も求められます。業務知識がなければデータもただの数字にすぎません。

●コミュニケーションスキル

データサイエンティストには、データ分析の必要性や分析結果を他の部門や経営陣が理解できるように、技術的な側面以外にも噛みくだいて説明したり、データに意味を持たせてプレゼンテーションする能力が求められます。また、現場の課題を汲み取る理解力も必要です。

そのため、他部門や経営陣との意思疎通を図るコミュニケーションスキルも求められます。

●直感・閃き、倫理的判断

データサイエンティストはIT技術を駆使してビッグデータを解析しますが、ビジネスの課題を解決したり事業戦略をサポートしたりするためには、まずデータにどのようにアプローチすれば良いのか、あるいはどのような仮説を立てるべきか、といった直感や閃きを必要とする、更には倫理的に善悪を判断するような、人間ならではの能力も必要とされます。

ここに、後述するAIとの共存の可能性があります。

データサイエンティストのスキルからプログラミングが消える？

先に、データサイエンティストにはITスキルとしてプログラミング言語の知識が必要であることを紹介しましたが、近年、データ分析のためのツールが進歩したことで、プログラミングを行わなくてもドラッグ・アンド・ドロップによるGUI上の操作でデータ分析を行えるようになってきました。

このことから、近々データサイエンティストの必須スキルからプログラミング能力の比重が軽くなっていくと思われれます。

データ分析のツールの例として、Microsoft社のAzure MLやGoogle社のCloud AutoMLなどがあります。

データサイエンティストとは何か。本当にAIに取って代わられるのか？

ただし、たとえドラッグ・アンド・ドロップで操作が行えるにしても、プログラミングの知識があるかないかでツールの使いこなしかたが変わってくるでしょう。

これは、WebデザインをGUIで行えるツールが登場しても、HTMLやJavaScript、PHPを理解しているデザイナーと理解していないデザイナーとでは、ツールの使いこなし方や仕上がったWebサイトの質に差が出ることと同じです。

データサイエンティスト自体が不要に？と考えるのはまだ早急

近年のAIの進歩は、さまざまな職種を人から奪うと言われていますが、データサイエンティストもその一つではないかと言われています。

これは、AIがビッグデータからパターンを抽出したり分類したりする能力に長けているためです。

実際、Amazonなどの Recommend 機能は、驚くほどの確に利用者に適した商品を推薦することがあります。

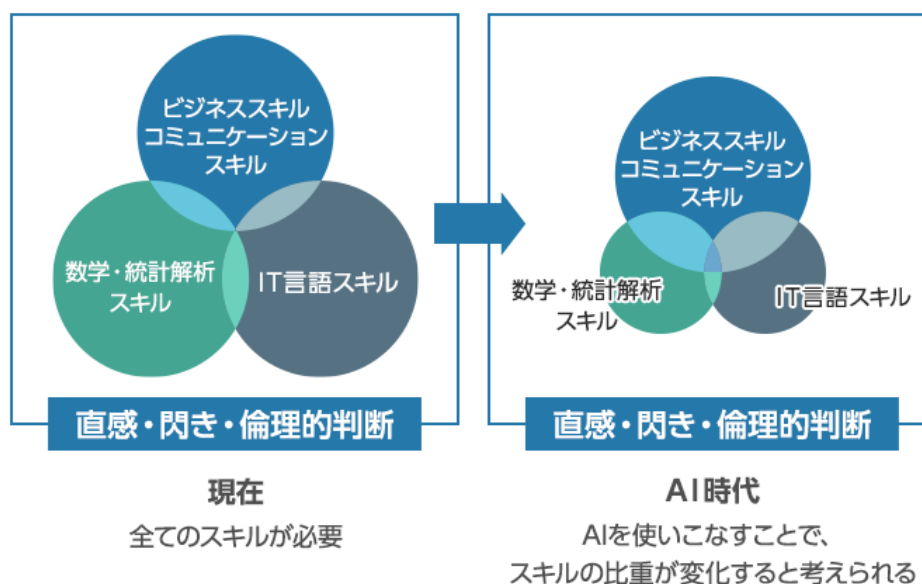
また、AIは機械学習によりビッグデータからレポートを作成するなどの作業を効率化します。

しかし、AIは過去のデータから一定のパターンや傾向を抽出することはできますが、人が働いている現場を観察したり現場の人たちに取材しながら課題を汲み取ったり、データに左右されない仮説を立てたり、あるいは事業戦略上のアイデアを立案することはできません。

したがって、データサイエンティストの業務はAIを使いこなすことで効率化することはできますが、データ分析を0から1にするにはまだ人の直感や閃き、あるいはマーケットセンスといった能力が必要とされています。

技術面でも、データをAIに処理させるまでの分析手法の組み合わせや、実際の処理のチューニングに人の手が必要な場面がまだまだ存在しています。

このことから、AIがデータサイエンティストの仕事をどこまで代行できるかどうかは、今の段階ではまだ分からないというのが現状です。



データサイエンティストとは何か。本当にAIに取って代わられるのか？

ニーズの高まりと競争の激化

ここまで見てきた通り、データサイエンティストはこれからも多くの企業から求められる人材です。

一方で、データサイエンティストに必要とされているスキルや能力については、明確な基準が存在しているわけではありません。また、AIの進歩やツールの高機能化により、データサイエンティストの業務範囲も変わってくるでしょう。

これらのことから、需要が高まっているデータサイエンティスト市場ですが、AI時代もさらに競争が厳しくなることが予想され、データサイエンティストをより適切に評価する仕組みが必要になっていくでしょう。

【制作／ブレイン】

※ Microsoft および Azure は、米国 Microsoft Corporation の、米国およびその他の国における登録商標です。

※ Google および Google Cloud Platform (GCP)は、Google LLCの商標または登録商標です。

※ 記載されている社名、商品名などは、各社の商標または登録商標である場合があります。

