

## 機械学習は「判断する機械」

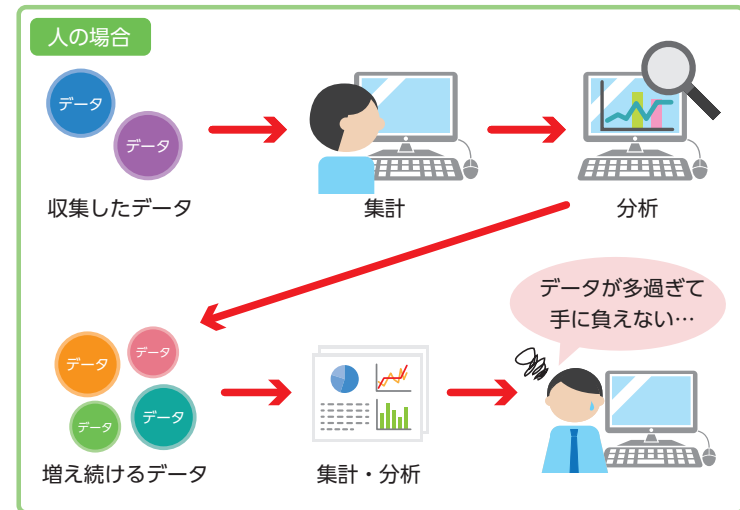
### コンピューター自身が判断するためのしくみ

現在、さまざまな産業分野で「**機械学習 (Machine Learning : マシンラーニング)**」が注目を集めています。機械学習が注目を集めるのは、「AI (人工知能)」や「自動運転」といった最新テクノロジーの開発や、既存のビジネスのデータ活用に、決して欠かせないしくみだからです。

機械学習とは、「**データの中から規則性や判断基準を見つけ、それを使って判断し、未知のものを予測する**」テクノロジーのことです。データから特徴や判断基準を見つけ出すことを、「**学習**」といいます。機械学習を活用することで、たとえば受信メールの中からスパムメールを見つけ出したり、たくさんの動物の写真の中から猫の写真だけを取り出して分類したりといった作業をコンピューターが自動でしてくれます。ただし、データおよび機械学習をどう活用するかは、最終的には人間が決定する必要があります。また、機械学習の精度を実際のビジネスに活用できるレベルまで高めるには、まだまだ人間のサポートが必要です。

世の中に蓄積された膨大な量のデータ、いわゆる「ビッグデータ」の分析にも機械学習の技術が欠かせません。機械学習のテクノロジーを駆使してコンピューターが処理すれば、膨大な量のデータを分析し、有益な情報を見つけ出すことができる可能性が上がります。人間や既存の統計ソフトウェアでは解決できないデータや、その他のデータを活用したい領域でも、機械学習なら通用することは多くあります。

### データから規則性を見つけ出し、自動で判断する



▲人間が与えた「答え」によって学習したり、膨大なデータの中から規則性や判断基準を見つけ出すことで、コンピューターが判断や予測を自動で行うことができる。

## 得意な分野と苦手な分野

### データに強いがデータがないと…?

機械学習が得意なことは、大きく3つに分けられます。1つ目は、**データをさまざまな角度から検証して、共通する特徴やパターンを見つけ出す**ことです。その中には、人間では思いつかないような特徴が見つかる可能性もあります。たとえばアメリカでは、がん患者の腫瘍の特徴を機械学習によって発見し、がんの診断に役立てています。

2つ目は**大量のデータを扱う**ことです。YouTubeの動画を1週間見せ続けたことにより、コンピューターが自分で「猫」の画像を認識できるようになったという実験があります。膨大な量のデータ（ビッグデータ）を有効活用できるということは、機械学習のビジネス利用において、もっとも期待されていることの1つです。

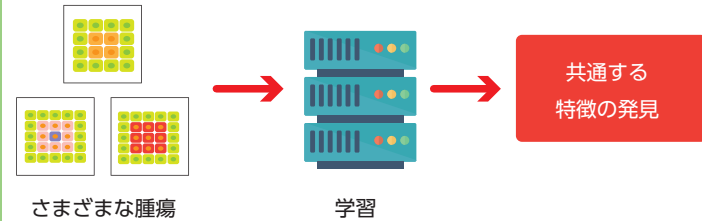
3つ目は、**決まったルールに則って、ブレずに判断する**ことです。機械は繰り返し処理が得意です。その点は根本的に機械が人間よりも優れている点だといえるでしょう。

反対に、「学習していないことには対応できない」という弱点もあります。強力な将棋ソフトが、初めて見る戦法に対応できずに敗れたことは有名な話です。わざわざコンピューターを使うまでもなく、人が作業したほうが早い規模の小さな仕事では、機械学習のメリットを活かせません。少ない情報から推論をすると判断の精度が上がらず、判断を誤ることもあります。

機械学習は、「データ」を「さまざまな角度から検証」し、「高速かつ正確に処理」する必要がある分野で、その力を最大限に発揮できます。

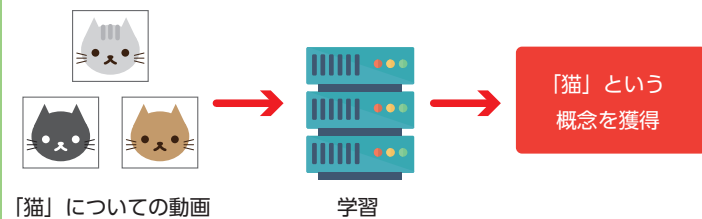
### 機械学習の得意なこと

#### 特徴の発見



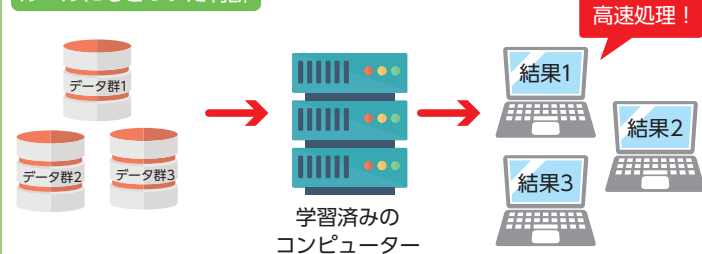
参考：IBM announces two new partnerships with MoleMap New Zealand and Melanoma Institute Australia to advance research in melanoma identification (<https://www-03.ibm.com/press/nz/en/pressrelease/50055.wss>)

#### 大量のデータ処理



参考：Using large-scale brain simulations for machine learning and A.I. (<https://googleblog.blogspot.jp/2012/06/using-large-scale-brain-simulations-for.html>)

#### ルールにもとづいた判断



▲機械学習は、多くの場合データが大量であればあるほど、その力を発揮してくれる。

## 機械学習とAIの違いって何？

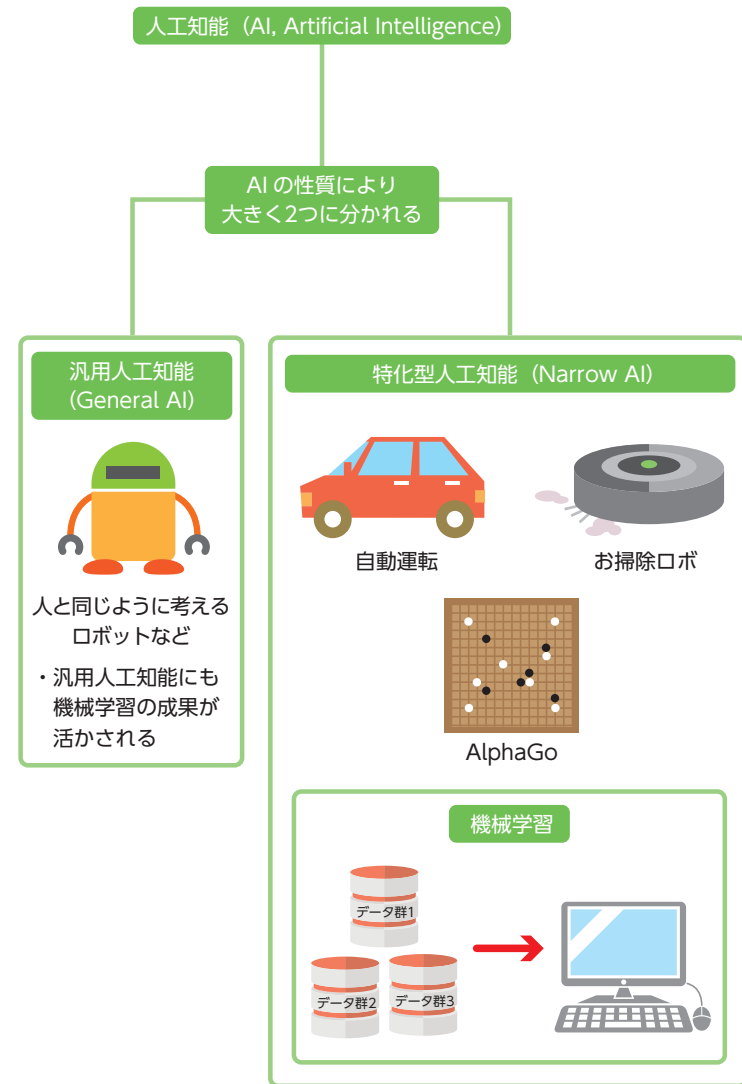
### 機械学習はAIの1分野

インターネットをはじめとするさまざまなメディアで「機械学習」が取り上げられるとき、同じように使われる言葉に「AI(人工知能)」があります。機械学習とAIの違いは何でしょうか。

AIは、大きく分けると「汎用人工知能 (General AI)」と「特化型人工知能 (Narrow AI)」の2つに分類することができます。汎用人工知能とは、かんたんにいうと「人間と同じように考え行動する」ことを再現しようというものです。一方で特化型人工知能は、人間が頭で行っている作業の一部をマシンに行わせようとするものです。汎用型人工知能は、たとえばアニメに出てくるアンドロイドやロボットをイメージしてもらえばいいです。また、最近話題になったソフトバンクの「Pepper (ペッパー)」も、この範囲に入ると考えることもできるでしょう。一方、特化型人工知能は、過去のデータの蓄積から学習し、特定の分野の問題解決や作業を独自に行うことができる人工知能です。現在活用されているAIや、近い将来製品化されるAIの多くが、この特化型人工知能に含まれます。

そして機械学習は、特定の事象のデータを解析して、その中から特徴やルールを学習し、判断や予測を行うテクノロジーであり、**特化型人工知能の1つのカテゴリ**であるといえます。そのため、機械学習はAIの1分野として研究されてきた歴史があります。同時に機械学習の発展がAI全体を押し上げ、ニューラルネットワークが発展してできたディープラーニングは、近年のAIブームの火付け役となりました。

### 機械学習は特化型人工知能の1分野



▲人工知能は、人間と同じように考え行動できる「汎用人工知能」と、特定の機能に特化した「特化型人工知能」に分かれ、機械学習は特化型人工知能の1分野であるといえる。

## 機械学習と データサイエンスの違いとは？

### 機械学習はデータサイエンスを支えるテクノロジー

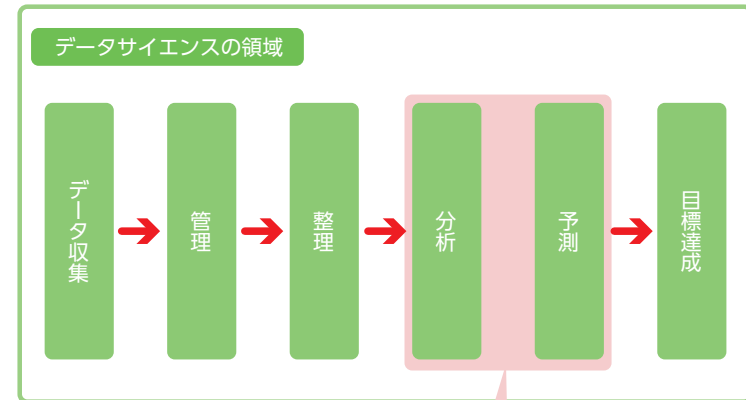
機械学習はデータから学習することで予測や判断を自動化する技術、対してデータサイエンスとはデータの分析や活用方法などを研究する分野です。どちらも「データをビジネスに活かす」ためのものですが、これらの違いはどこにあるのでしょうか。機械学習はデータサイエンスの一部という見方が有力です。

企業の中には、膨大なデータが蓄積されています。それらのデータを有効に活用し、企業戦略の意思決定をはじめ、さまざまな業務で活用したいと考えるのは、当然のことでしょう。

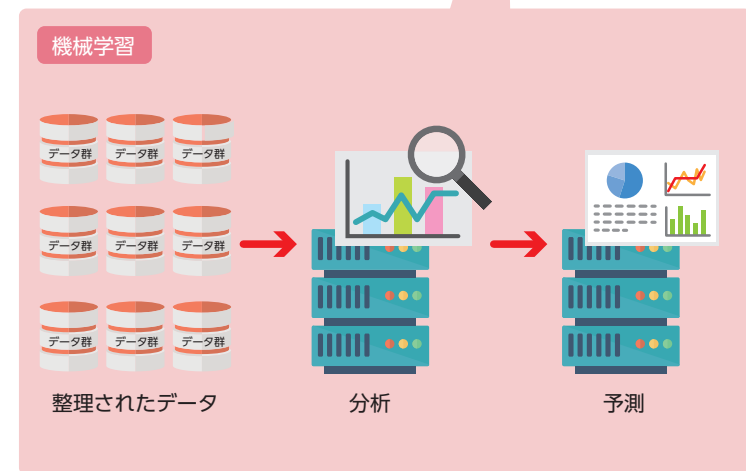
しかし、データを正しく活用するためには、「データを集める」「管理する」「整理する」「分析する」「予測する」といった、複数の工程が必要になります。このデータ活用の範囲をすべてカバーするのが、**データサイエンス**です。このようなデータ活用のための諸業務を行う人のことを、「**データサイエンティスト**」や「**データアナリスト**」と呼びます。これらの作業はいずれも高い専門性が求められます。

データ活用全体の視点に立つデータサイエンスに対して、機械学習は分析・予測などデータ活用の一部に特化しています。機械学習を活用するためには、その前段としてデータの収集・整理も必要なので、データサイエンス全体の知識は必須です。**機械学習がデータを有効利用するための一手段であるのに対し、データサイエンスはそれを包括した、データ活用の総合的な分野**でお互い欠かせないものといえるでしょう。

### データサイエンスに機械学習は欠かせない



データサイエンスの一部に機械学習が含まれるという見方が多い



機械学習は、今日のデータ分析において欠かせないテクノロジーであり、もっとも重要な手法の1つとなりつつある

▲データサイエンスの「分析」と「予測」の段階で、機械学習の手法が必要となる。