

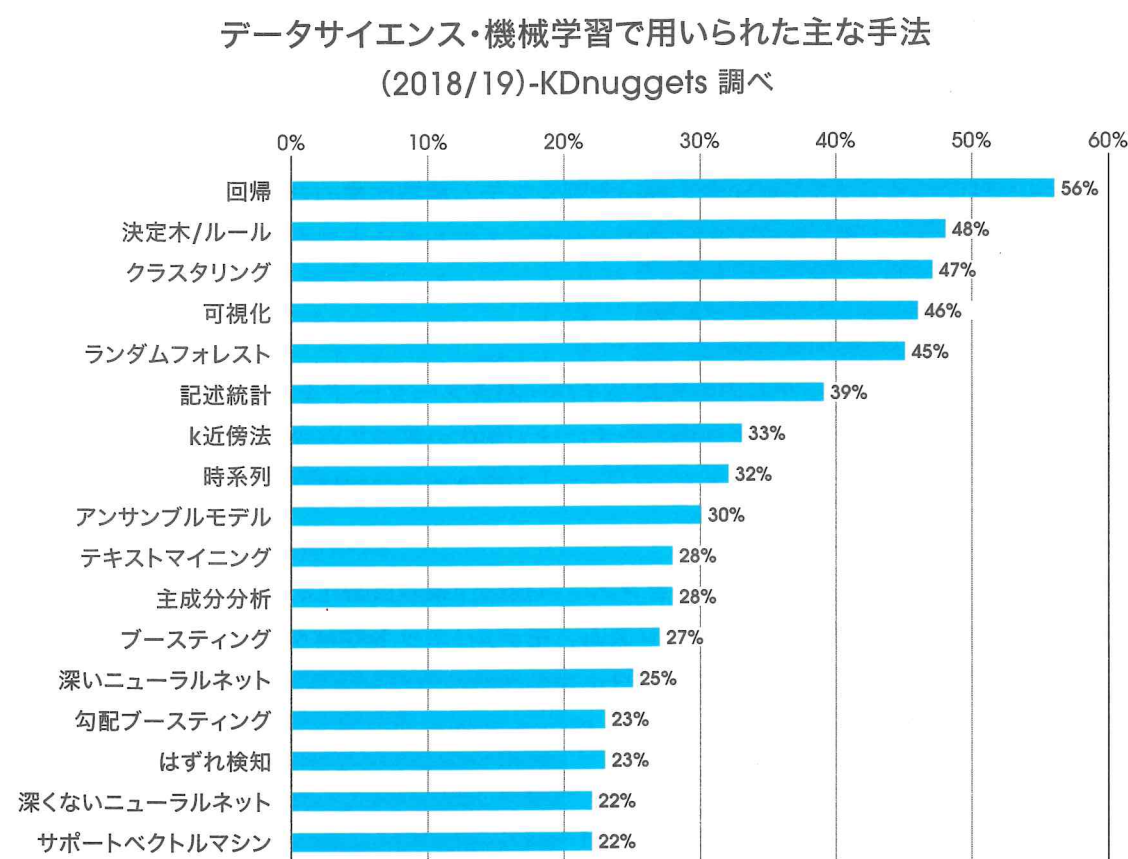
### 機械学習・データサイエンスの業務従事者サーベイにみる「学習」のトレンド

Kaggleは、世界中の企業や研究者がデータやモデルを投稿するプラットフォームで、世界中の機械学習・データサイエンスに携わっている約60万人が集まるコミュニティを形成している。

「Competition (コンペ)」と呼ばれる、企業や政府が課題を提示し、参加者がオープンな場で競争を行い、賞金と引き換えに企業がモデルを買い取るという仕組みを提供している。Kaggleでは毎年、自社プラットフォームの利用者サーベイを行っており、このデータを、データ分析及び可視化のためのオープンデータとして公開している。2018年版のサーベイデータ<sup>※22</sup>によると、最も使われている言語はPythonであり、最も利用されているフレームワークはScikit-LearnやTensorFlow(ディープラーニングのフレームワーク)である。

KDNuggetsによるサーベイ結果<sup>※23</sup>では、データサイエンス及び機械学習分野で、よく利用されているアルゴリズムとしては、回帰(Regression)、決定木(Decision Trees/Rules)、クラスタリングが上位を占める(図2-2-11)。

■図2-2-11 KDNuggetsのサーベイ(2018/19)<sup>※23</sup>



出典: "Top Data Science, Machine Learning Methods Used, 2018/2019" (KDNuggets.com)

#### 参考文献

- [1] David H. Wolpert and William G. Macready, "No Free Lunch Theorems for Search", Technical Report SFI-TR-95-02-010, Santa Fe Institute, 1995.
- [2] Satoshi Watanabe (1969). Knowing and Guessing: A Quantitative Study of Inference and Information. New York: Wiley. ISBN 0-471-92130-0. LCCN 68-56165.

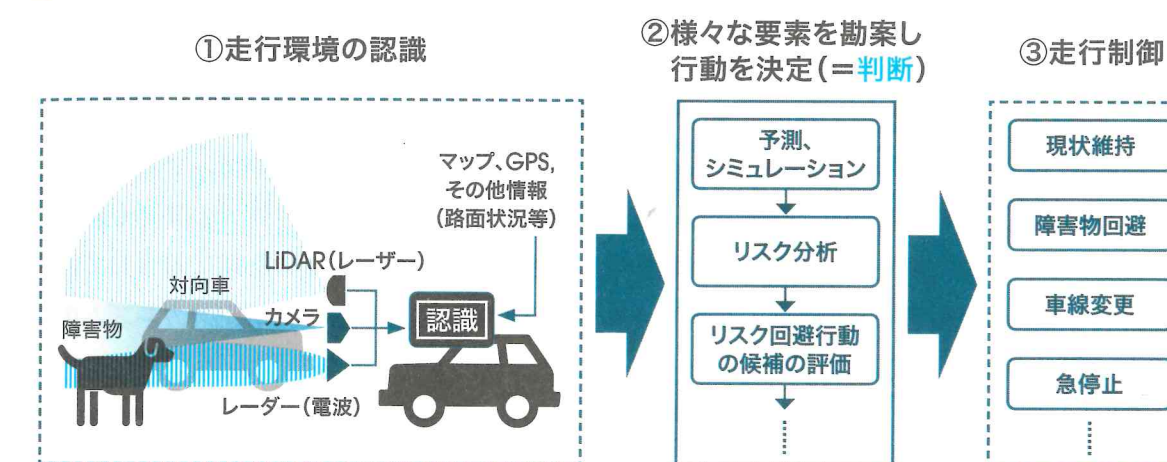
※22 <https://www.kaggle.com/paultimothymooney/2018-kaggle-machine-learning-data-science-survey>

※23 <https://www.kdnuggets.com/2019/04/top-data-science-machine-learning-methods-2018-2019.html>

### 2.2.4 判断

意思を決定するためには一般に複数の要素を勘案する必要がある。例えば自動運転車でも、「ネコが飛び出してきた」からハンドルを切るのではなく、「路面が滑りやすい」、「対向車が来ている」などそのほかの状況を合わせて総合的に「判断」する必要がある(図2-2-12)。「判断」については、後述のルールベースなど従来のAI技術も有効であるが、AlphaGoで注目を浴びたように機械学習による技術も登場している。本節では、AIによる「判断」に焦点を当て、データを解析した結果に基づく判断=意思決定する技術、及び、自動交渉、言論マップ、議論マイニング、ディベートAIなどの最新技術動向を紹介する。

■図2-2-12 自動運転における「判断」の例



#### (1) 概要

ここでは、「判断」とは、複雑な制約がある中で、何らかの評価関数を最大にするように、意思決定をすることとする。意思決定の対象は、単独の行動や、組み合わせ(マッチング)や順番の決定(プランニング)である。

「意思決定」は、睡眠中を除けば、常に行われる活動の一つだ。この仕事を誰に頼むか、昼食はどこで食べるか、気に入った商品を注文するかなど、様々な局面で意思決定を行う。しかし、意思決定の誤りが業務上、重大な結果を引き起こすことがあれば、インターネットを通じて大量の情報が流れ込み、判断に迷うこともある。こうした状況で、AIに判断材料を提供させ、「意思決定」の参考とすることは有用である。また自動運転車やロボットのように、AI自身に「判断」を行わせるケースもある。

AIが判断を行う手法としては、第二次AIブームのときに注目されたルールベースがある。人間が意思決定する際には、法律や企業の内規、経験や嗜好からつくられた「こういう場合にはこうすべき」という自分なりのルールを勘案するであろう。このようなルールを「if...then...」形式で記述し、確からしさを表す数値を付与してデータベースに蓄積したものをルールベースと呼ぶ。実行時には、現在の状況に関係する全てのルールを評価し、ifの部分が一致した場合にthenに書かれた処理を実行し、確からしさを表す数値を積算して総合的な確からしさを判定する。

別の手法として、決定木分析がある。データを性質(例:男女)や実績値(例:販売数5個以上/以下)に基づき木構造で分析しておくことで、現状と照らし合わせて判断を行わせることができる。また、将棋や碁のように、現在の状況で打てる手の中でより効果が高い一手を探索していくこともできる。