

AIでセレンディピティや多様性を生み出す

◆セレンディピティ（予想外の発見）を生むAIによる材料開発

2020年5月、理化学研究所や物質・材料研究機構などの研究チームは、材料開発において、既知の傾向から外れる例外的な特性を持つ物質の発見に特化した人工知能（AI）、「BLOX」を開発し、新規の光吸収特性を持つ物質を発見した。

一般的なAIを活用した材料開発では、既存のデータセットを学習させたモデルから、目標とした特性を持つ物質を探索する方法が多くある。一方で、この方法は予測する特性を限定しており、学習に用いたデータの傾向から外れる、例外的な特性を持つ物質の発見には適していない。目標値を設定せずに、ランダムに傾向から外れた特性の物質を探索するAIもあるが、やはり探索する範囲を限定する必要がある、優れた特性の物質を見落とす可能性や、探索に時間がかかるなどの課題がある。

BLOXは、まず一般的な機械学習法で複数の物質について特性を予測し、次に、予測された特性の分布や傾向から外れた、例外的な特性を持つ物質を抽出する。さらに、抽出した物質の情報を既知のデータとして反映し、以降に行われる物質の特性予測に用いる。材料特性の予測と、傾向から外れた例外的な特性を持つ物質の抽出・反映を繰り返すことで、特性の分布を拡大させ、従来知られている傾向からは予想しにくい物質を効率的に見つけることができる。

研究チームはこのAIを活用し、光吸収強度と吸収波長の2つの特性に対して、例外的な特性を持つ有機化合物の候補を、従来の方法よりも効率的に多数発見した。また、発見した候補の中から8個を実際の化合物で評価した結果、250ナノメートル以下や450ナノメートル以上の領域で、光吸収強度が高い化合物が存在することを確認した。

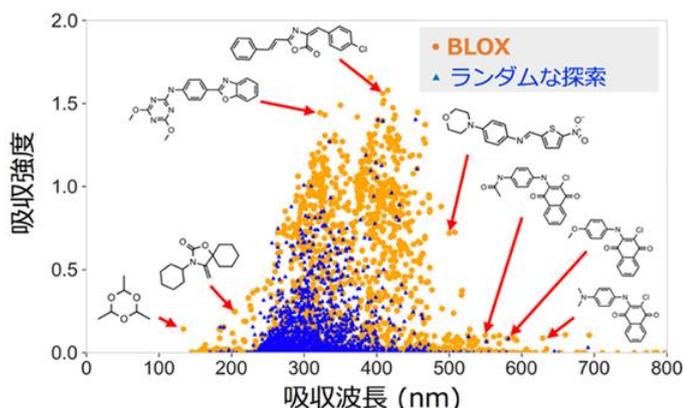


図 例外的な光吸収特性をもつ化合物の探索結果
出所) <https://www.nims.go.jp/news/press/2020/05/202005281.html>

◆好奇心のあるAIがデータの多様性を拡大する

20年1月、グラスゴー大学の研究チームは、自動実験装置に、「好奇心アルゴリズム」と呼ばれる、人間のような好奇心に基づく学習方法を取り入れ、効率的に多様な実験結果を得ることに成功した。

一般的に、多様な実験結果を得る方法として、統計的手法やランダム化によって、実験の条件を分散させるものが知られている。一方、好奇心アルゴリズムは、これまでに取得した実験結果の分布を踏まえて、データが存在しない未知の領域に仮の目標値を設定することが特徴である。未知の領域に目標を設定することを、人間の好奇心になぞらえており、目標設定と実験を繰り返す。

研究チームは、液体中の油滴挙動を観察する実験を、一般的なランダムに実験条件を変更する探索手法と、好奇心アルゴリズムで比較して行った。前者では、条件を変えながら1000回の実験を行ったが、温度条件と液滴挙動の有意な関係性は観測できなかった。一方、後者では、300回程度の試行で、温度と液滴挙動の強い相関性を示すデータが得られた。実験条件をランダムに変更するより、目標値を仮定して未知の領域を探索する好奇心アルゴリズムの方が、より効率的に多様なデータを収集できる。

◆AIに利用するデータにも多様性が必要

AIの活用においては、データの量や偏りが予測の質を左右するため、学習に用いるデータの多様性も求められる。例えばモビリティ分野では、自動運転の正確な判断のために、事故につながるような極めて特殊な走行シーンのデータも要求される。金融分野では、学習データの偏りによって、AI審査が倫理的に誤った判断を行うことを避けるため、性別や人種など多様性のあるデータが求められる。

20年2月に総務省が発表した「AI経済に関する現状等」によると、AIに活用されるデータの価値は、量、精度、鮮度、多様性などで変動する。また「AIの利用などを巡る産業などの構造」では、国内のAI利用における課題として、データが単一企業でのみ使用される傾向にあり、多様性に欠ける点が指摘されている。

さまざまな分野でAIが利用され、存在感を増しているが、既存のデータと特定の目標に限定した予測だけではAIの活用は頭打ちになる。AIでデータの多様性やセレンディピティを生み出す取り組みの重要性が高まっている。 【塚原祐介】