

疑似量子の活用で従来の MI による材料開発期間をさらに 20%短縮できることを

実証

疑似量子コンピュータ技術を適用した機械学習により性能予測をさらに高精度化

株式会社日立製作所(以下、日立)は、材料開発の加速につながる新たな機械学習モデルを開発し、積水化学工業株式会社と進めているマテリアルズ・インフォマティクス(以下、MI)推進に向けた協創活動^{*1}においてその有効性を実証しました。本モデルは、決定木を用いた材料の性能予測モデルの構築に、量子コンピュータを疑似的に再現する「CMOS アニーリング^{*2}」を適用することで、さまざまな条件を網羅的に考慮することが可能です。これにより従来の機械学習モデルの予測精度を本モデルの適用により向上させ、材料開発の期間を約 20%短縮できる見通しを得ました^{*3}(図 1)。日立は、本技術を Lumada の「材料開発ソリューション^{*4}」として実用化していきます。また、本モデルを生産現場の運用や物流の最適化など、幅広い業界への展開をめざし、お客さまの価値創造に貢献していきます。

なお、本技術の一部は、12月17日から20日に開催される IEEE Bigdata 2022 において発表予定です。

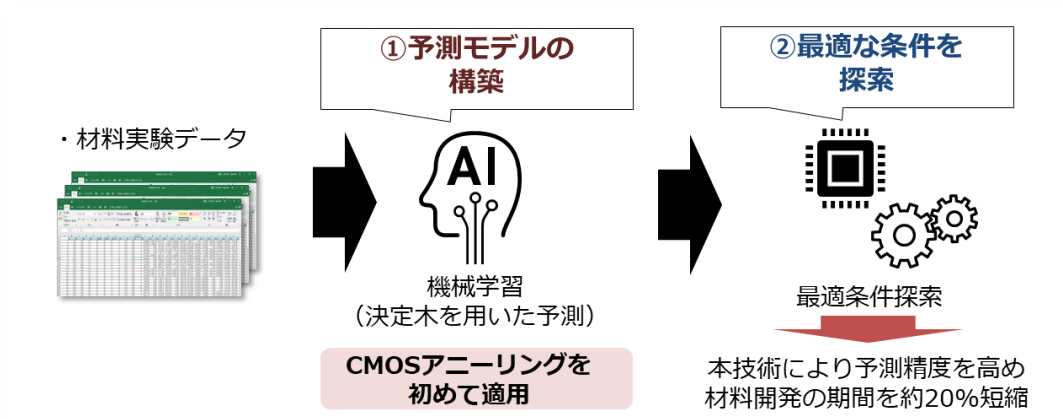


図 1 開発した機械学習モデルの概念図

これまでの材料開発は研究者の経験や勘に依存し、実用化までに長い時間と費用を要していました。一方で、市場では製品ライフサイクルの短命化、資源の制約などを起点として材料開発への要求が多様化・高度化しており、その要求に応えるべく MI への取り組みが本格化しています。日立はデータを活用して材料開発プロセスを効率化するデータ管理・アナリティクス基盤の実現に向けて、技術開発を積み重ねるとともに MI を適用した「材料開発ソリューション」を Lumada ソリューションの 1 つとして提供し、素材産業を始めとするさまざまなお客さまのニーズに応えてきました。

新たな材料の開発では、予め機械学習により材料の特性を予測し、有望な材料に絞って開発を進める

ことで開発期間を短縮する取り組みが検討されています。具体的には、①材料に関する既存のデータを AI が学習することにより、材料の特性を予測する予測モデルを構築、②得られた予測モデルを用いて、実際に開発したい材料の作成に最適な条件を探索します。このような取り組みで用いられる機械学習には、少ない実験で有用な材料の候補を見つけるために、材料の特性を高い精度で予測する技術が求められます。予測精度の向上には AI が複数の変数を組み合わせた複合条件で学習することが有効ですが、材料の作成に関わるさまざまな条件を考慮するとその組み合わせは膨大になります。従来はデータサイエンティストが経験や勘に基づき組み合わせる条件を選び、予測モデルを構築していましたが、条件を網羅的に検討していないため精度向上に限界がありました。

今回、日立は、機械学習モデルの構築(上述の①)において、MI でも一般的に用いられる決定木アルゴリズム^{*5}に、疑似量子コンピュータ技術である CMOS アニールングを導入し、材料の作成に関わる条件を網羅的に考慮することで予測精度を高める新たな手法を開発しました。このようにして構築した高精度予測モデルを、既存の有機材料の分子構造からその特性を予測する問題に適用した結果(上述の②)、図 2 に示すとおり、機械学習による予測精度向上において高い性能を示す従来技術 LightGBM、および XGBoost^{*6}に本手法を適用した際の予測精度が、従来技術単独での予測を上回る精度^{*7}を達成しました。さらに、本技術を用いて材料開発を行った場合、従来の MI を用いた方式に比べ、開発に要する期間を約 20%削減できる見通しを得ました。

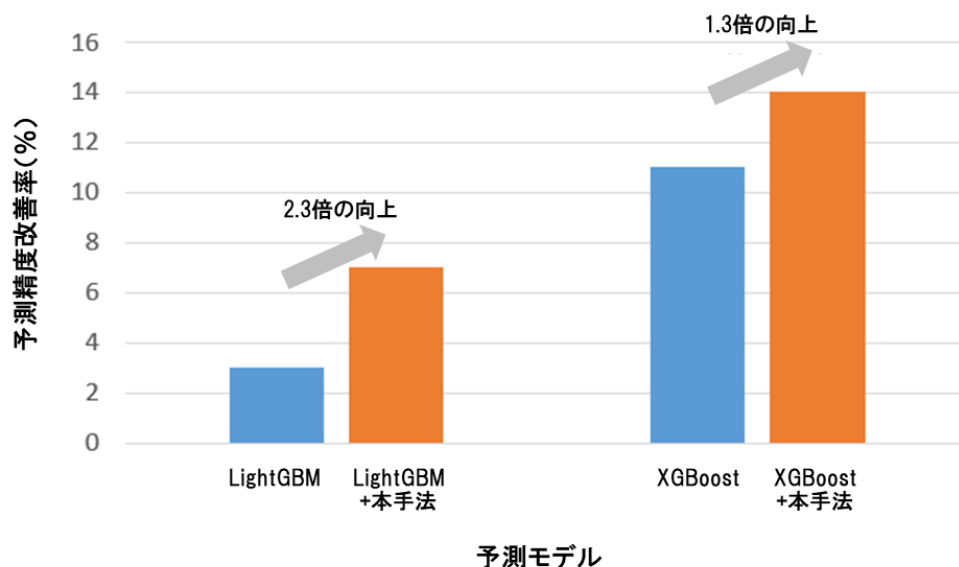


図 2 従来技術に開発した手法を適用した際の予測精度の改善結果

本技術は、決定木アルゴリズムを用いる機械学習の予測精度向上に広範に適用可能です。今後、Lumada の「材料開発ソリューション」として実用化を進め、半導体や電池、創薬など、お客さまの多様な材料開発の効率化をめざすとともに、物流、生産領域含めた分野でも本技術の適用を推進することで、お客さまの創出する社会価値向上の加速と、サステナブルな社会の実現に貢献していきます。

■補足説明：開発した技術の詳細

1. 複合条件を用いてデータを分類する決定木アルゴリズム

決定木は入力データを条件に基づき次々分類していく手法で、その予測精度は分類に用いる条件や、条件の組合せにより決まります。一般の計算機を用いる場合、決定木アルゴリズムは、計算が膨大とならないように一つの変数を条件に用いてデータを分類していました。ところが、材料開発のようにさまざまな変数が材料の特性に影響する問題を予測するには、決定木の扱う条件が単純すぎるという問題がありました。そこで、複数の変数が組み合わされた複合条件を論理積で表現することで、多数の変数から自動かつ直接的に生成する技術を開発しました(図 3)。これにより、例えば、アイスクリームの売上を予測する場合、従来は、晴れか否かで分類し、次に午後か否かで分類、というように一つの変数で分類していましたが、本技術により、晴れて気温 25 度以上で湿度 40%以上か否かで分類、次に、土曜日の午後か否かで分類、というように複数の変数が複合した条件を用いてデータを分類することが可能になりました。

2. 決定木アルゴリズムの複合条件を最適化する CMOS アニーリング技術

上記決定木アルゴリズムでは、例えば、100 個の変数の中から 3 つを組み合わせることでデータを分類しますが、その組み合わせ総数は 10 万通り以上になります。予測精度が最高になる組合せを求めると、全ての組合せに対して計算すると膨大な時間を必要とします。このような多くの組合せの中から最適なものを短時間で選び出す問題には疑似量子コンピュータ技術である CMOS アニーリングが適しています。そこで、決定木アルゴリズムで最適な複合条件を選び出す問題を、CMOS アニーリングが扱える QUBO 形式の問題⁸に変換して計算した結果、現実的な時間内で、予測精度が最も高まる複合条件を見出すことが可能になりました。データサイエンティストは経験や勘に基づき複合条件を決定しますが、本技術は全ての組合せの中から最適なものを自動かつ直接的に選べる特長があります。本技術は、機械学習と CMOS アニーリングという異なる分野の研究者の知の融合により開発されました。

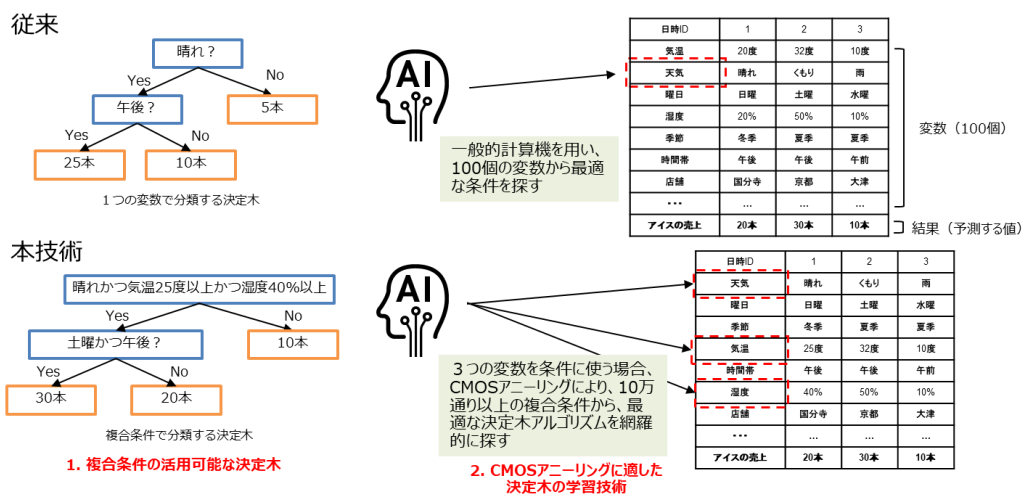


図 3 複合条件を用いてデータを分類する決定木アルゴリズム、および、複合条件を最適化する CMOS アニーリング技術

- *1 2022 年 9 月 20 日付ニュースリリース、「積水化学と日立、材料開発における MI 推進に向け、先進デジタル技術を用いた協創を開始」
<https://www.hitachi.co.jp/New/cnews/month/2022/09/0920.html>
- *2 [CMOS アニリング](#)：磁性体の性質を説明するために考案されたイジングモデルを用いて組合せ最適化問題を解くために日立が開発している新型コンピュータ。
- *3 予測された素材開発データの性能検証にかかる実験の回数を従来技術と本技術適用済の従来技術とで比較した結果。具体的には、過去の 100 個の素材開発データを元に、従来技術と本技術適用済の従来技術それぞれで性能予測を実施した結果、実際に評価の高い 10 個のデータを予測するまでに従来技術の場合は 24 個、本技術では 19 個の予測が必要だったことから、従来と比較し検証にかかる実験期間が 20%低減すると試算。
- *4 [材料開発ソリューション](#)：MI によるお客さまの新材料の開発を支援するサービス。
- *5 決定木アルゴリズム：分類問題と回帰問題を解く教師あり学習のアルゴリズムの一つ。与えられたデータに対し次々に条件を設けて、データを段階的に分類していく手法。
- *6 TreeBoosting-03: Why Does XGBoost Win Every Machine Learning Competition?
<https://datasciblog.github.io/2020/02/26/tree-boosting-03/>
- *7 ランダムフォレスト(2001)からの誤差改善率により評価。
<https://link.springer.com/article/10.1023/a:1010933404324>
- *8 QUBO：Quadratic Unconstrained Binary Optimization (二次制約なし二値最適化)。量子アニリングマシンなどのコンピュータで「組合せ最適化問題」を解く際に必要となる専門的な形式。具体的な数式は、以下となる。

$$\text{SWMSE} = \sum_{b=0,1} \text{Var} \left(\left\{ t^{(i)} \mid i \in \mathbb{S}_b \right\} \right) \left(\frac{N_{\mathbb{S}_b}}{N_{\mathbb{S}}} \right)^2$$

■日立製作所について

日立は、データとテクノロジーでサステナブルな社会を実現する社会イノベーション事業を推進しています。金融・官公庁・自治体・通信向け IT サービスやお客さまの DX を支援する「デジタルシステム&サービス」、エネルギーや鉄道で脱炭素社会の実現に貢献する「グリーンエナジー&モビリティ」、産業流通、水インフラ、ヘルスケア、家電・空調システム、計測分析システム、ビルシステムなどの幅広い領域でプロダクトをデジタルでつなぐ「コネクティブインダストリーズ」と、自動車・二輪車の分野で先進技術を提供する「オートモティブシステム」の事業体制のもと、IT や OT(制御・運用技術)、プロダクトを活用する Lumada ソリューションを通じてお客さまや社会の課題を解決します。グリーン、デジタル、イノベーションを原動力に、お客さまとの協創で成長をめざします。2021 年度(2022 年 3 月期)の連結売上収益は 10 兆 2,646 億円、2022 年 3 月末時点で連結子会社は 853 社、全世界で約 37 万人の従業員を擁しています。

詳しくは、日立のウェブサイト(<https://www.hitachi.co.jp/>)をご覧ください。

■お問い合わせ先

株式会社日立製作所 研究開発グループ

問い合わせフォーム：<https://www8.hitachi.co.jp/inquiry/hqrd/news/jp/form.jsp>

以上

このニュースリリース記載の情報(製品価格、製品仕様、サービスの内容、発売日、お問い合わせ先、URL 等)は、発表日現在の情報です。予告なしに変更され、検索日と情報が異なる可能性もありますので、あらかじめご了承ください。
