

## 最少の実験回数で高い予測精度を与える汎用的 AI 技術を開発

～材料開発の DX : NIMS、旭化成、三菱ケミカル、三井化学、住友化学の水平連携で実現～

国立研究開発法人物質・材料研究機構

旭化成株式会社

三菱ケミカル株式会社

三井化学株式会社

住友化学株式会社

### 概要

- 国立研究開発法人物質・材料研究機構 (NIMS)、旭化成株式会社、三菱ケミカル株式会社、三井化学株式会社、住友化学株式会社は、化学マテリアルズオープンプラットフォーム (化学 MOP) からなる水平連携において、強度や脆さといった材料物性を機械学習で予測する際に、材料の構造から得られる情報を有効に活用し、少ない実験回数で、予測値と実値の誤差を小さくできる (予測精度の高い) AI 技術を開発しました。高分子材料をはじめとした材料開発の強力なツールになると期待されます。
- これまでのマテリアルズ・インフォマティクス研究は、材料組成や加工プロセス (温度や圧力など) のパラメータから材料物性を機械学習で予測することで、材料開発を加速してきました。一方で、プロセス加工後の構造が材料物性に強く影響する場合、高い予測精度を実現するためには、構造情報を提供する X 線回折 (XRD) や示差走査熱量測定 (DSC) 等の測定データの利用が有効です。しかし、これらの測定データは、プロセス加工した材料に対して測定しないと取得できません。したがって、構造情報を利用して予測精度を向上させるには、研究者が設定可能な材料組成といったパラメータと、実測でしか得られないパラメータの異なる 2 つのパラメータを扱う必要があります。
- 本研究では、XRD や DSC 等の実測でしか得られないデータを用い、なるべく少ない材料作製回数で正確に材料物性が予測できるように、作製すべき材料を適切に選定する AI 技術を開発しました。作製すべき材料をベイズ最適化などの手法で選定し、測定したデータを加えて、AI による材料選定を繰り返します。本技術の有用性を示すための一例として、高分子材料であるポリオレフィンのデータベースを利用しました。その結果、AI 技術の利用により、無作為に材料作製を進める場合と比べて、作製回数を少なくしても機械学習による材料物性の予測精度を向上できることを示しました。
- 本技術を利用して精度の高い予測が実現できると、材料の「構造」と「物性」の関係が明らかになり、物性の発現起源の明確化・材料開発指針の決定が可能となります。また、この技術はポリオレフィンといった高分子材料だけでなく様々な材料開発にも応用できる汎用的技術です。そのため、材料開発の DX 基盤技術になると期待しています。
- 本研究は、NIMS 田村亮主任研究員、中西尚志グループリーダー、出村雅彦部門長、旭化成株式会社 武井祐樹主幹研究員、三菱ケミカル株式会社 今井真一郎主席研究員、三井化学株式会社 中原真希研究員、住友化学株式会社 柴田悟史研究員によって実施されました。
- 本研究成果は、Science and Technology of Advanced Materials: Methods 誌に、2021 年 9 月 28 日 (日本時間) にオンライン掲載されました。

### 材料物性予測精度を向上させる AI 技術開発

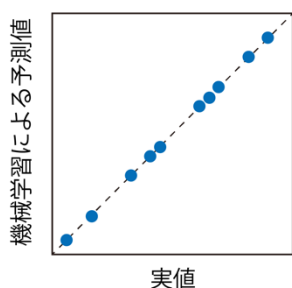


## 研究の背景

これまで一般的な材料開発では、作製する材料を専門家が長年の知識・ノウハウによって選定することで優れた材料が開発されてきました。この専門家による選定を AI に置き換えることでさらに材料開発を加速したいというニーズの元、マテリアルズ・インフォマティクス (MI) 研究が活発に実施されています。近年、材料物性の向上を目指すために、作製すべき材料を適切に選定できる AI 技術が、ベイズ最適化<sup>1</sup>などを利用し数多く開発されています。

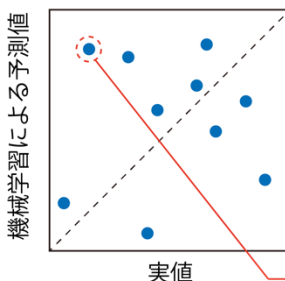
一方で、機械学習による材料物性を予測する精度を高くすることは、MI 分野において効率の良い材料開発の実現および材料を詳しく理解するために不可欠です。機械学習による予測が正確であれば、材料物性に関する測定実験をしなくても、材料物性を予知することができます (図 1 左)。一方で、予測精度が低い場合、予測値と実値の誤差が大きく (図 1 右)、これでは材料開発を加速することはできません。そのため、できるだけ少ない材料作製回数で、機械学習の予測値と実値の誤差をより小さくする (より高い予測精度を与える) 技術開発が重要となります。これを実現するために、作製すべき材料を適切に提案できる AI 技術の開発が始まっています。

### 予測精度が高い



➡ 予測値と実値の誤差が小さい

### 予測精度が低い



➡ 予測値と実値の誤差が大きい

予測値が大きかったのに  
作製したら、実値は小さかった

図 1 機械学習の予測精度が高い場合 (左図) と低い場合 (右図) の例。予測精度が高い場合、機械学習による予測値と実値の誤差が小さいため、材料物性を正確に予知することができます。

## 研究内容と成果

材料を表現するパラメータである記述子<sup>2</sup>には、研究者が設定できるものと容易にはできないものがあります。前者は、材料組成や加工プロセス (温度や圧力など) のパラメータです。後者の研究者が容易には設定できないが材料を記述する上で重要な記述子として、プロセス加工後の材料の構造情報を提供する XRD や DSC 等の測定から得られるデータが挙げられます。例えば、高分子材料においては、この記述子は高次構造<sup>3</sup>の情報を含んでいます。

これまでの MI 研究は、主に前者の材料組成や加工プロセス (温度や圧力など) のパラメータから材料物性を機械学習で予測することで、材料開発を加速してきました。一方で、プロセス加工後の構造が材料物性に強く影響する場合、より高い予測精度を実現するには、後者の記述子にあたる、XRD や DSC 等の測定データの利用が有効です。では、測定で得られたデータから材料物性予測のための機械学習の精度を向上させるためには、どのような材料を作製すれば良いのでしょうか? 機械学習の予測精度を向上させる手法として、能動学習 (アクティブラーニング) の一種である不確実性サンプリング (uncertainty sampling)<sup>4</sup>が有用です。しかし、この手法をそのまま利用すると、「ある特定の XRD パターンや DSC プロファイルを持つ材料を作ることによって機械学習の予測精度を高くすることができるため、その材料を作製してください」と提案されます。果たして、「指定されたデータに合致する材料」を作ることができるのでしょうか? XRD パターンや DSC プロファイルは、材料をプロセス加工し、測定を行って初めてわかるものであり、それを持つ材料をピンポイントで作製することは極めて困難です。そのため、研究者が設定できるパラメータと予測に利用する実測でしか得られないパラメータが異なる場合にも対応できる新しい手法を開発する必要

がありました。本研究では、XRD や DSC 等のデータを用い、なるべく少ない材料作製回数で、正確に材料物性が予測できるように、作製すべき材料として研究者が設定できる記述子を適切に選定する AI 技術を開発しました。(図2)。

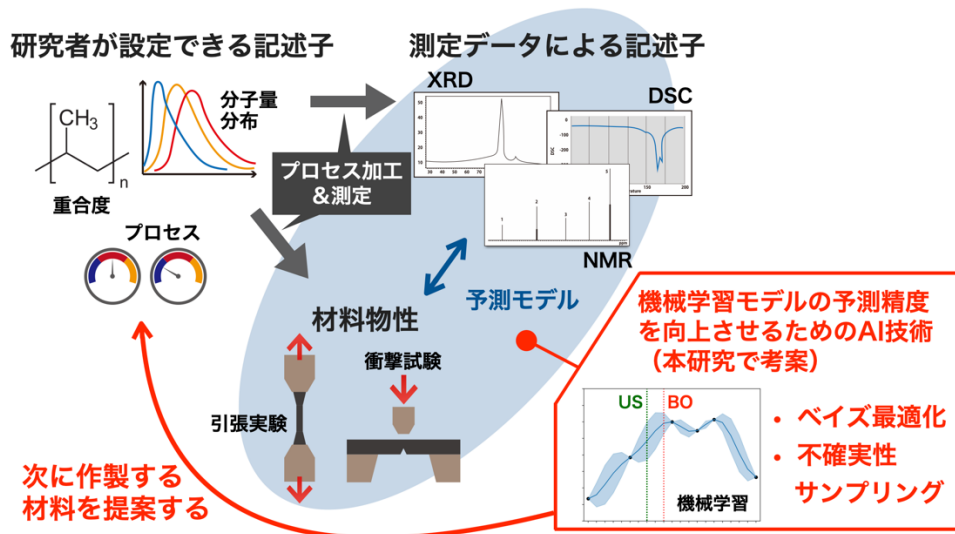


図2 研究者が設定できる記述子（組成や加工プロセスなど）とプロセス加工後の材料の構造を提供する XRD や DSC 等の測定データによる記述子、および材料物性の関係。本研究で考案した AI 技術は、測定データによる記述子から材料物性を予測する機械学習の精度を高くするための方法です。AI 技術を用いて、研究者が設定できる記述子を適切に選択し、実際に材料を作製することで学習データを増やし、予測精度を向上させます。

開発 AI 技術では第一に、研究者が設定できる記述子のみで構成された候補材料データを用意します。この候補データが1つ選択されると、その情報を用いて材料を1つ作製することができます。材料が作製されると、プロセス加工後の材料に対する XRD や DSC 等の測定データと材料物性を得ることができます。第二に、初期データとして、いくつかの候補材料に対して、XRD や DSC 等の測定データと材料物性を全て含んだデータセットを準備します。このデータセットを用いて、測定データから材料物性を予測する機械学習モデルを構築します。第三に、この機械学習の予測精度が高くなると期待される候補材料を AI が適切に選択し、実際に研究者が材料を作製します。この AI による選択の方法として、ベイズ最適化に基づく手法 (BOED) と、不確実性サンプリングに基づく手法 (USED) の2種類を提案しました。この AI による選択・実験を繰り返すことにより、少ない実験回数で、予測精度の向上を目指します。

この提案した手法の有効性を検証するために、一例としてポリオレフィンを対象材料とし、化学 MOP において作成したデータベースを利用しました。このデータベースは、15 種類のポリプロピレンについて、各々5つの異なるプロセス加工により得られた合計 75 種類のサンプルによって構成されています。対象とする材料物性は、シャルピー衝撃試験と引張弾性率の機械物性とししました。研究者が設定できる記述子として、分子量、立体規則性、および射出成型冷却温度に設定しました。また、測定によって得られる記述子としては、XRD や DSC 等を対象とししました。図3に予測誤差の材料作製回数依存性を示しました。予測誤差はクロスバリデーション<sup>5</sup>により評価しており、図3縦軸の値が小さければ小さいほど、機械学習の予測精度が高いことを意味しています(値が小さい方が図1左に対応)。初期データ数として、10 個の無作為に選んだ候補データからスタートしました。AI の支援を受けずに実験を進める場合(無作為に材料を作製する場合: 橙色のマーカー)と比べて、提案した2つの AI 技術を利用することで、シャルピー衝撃試験と引張弾性率のいずれの機械物性についても少ない材料作製回数で図3縦軸の値(予測誤差)が小さくなっていることがわかります。つまり、本研究で提案した BOED と USED のどちらの AI 技術を利用しても、少ない材料作製回数でより高精度な機械学習が実現できることがわかりました。

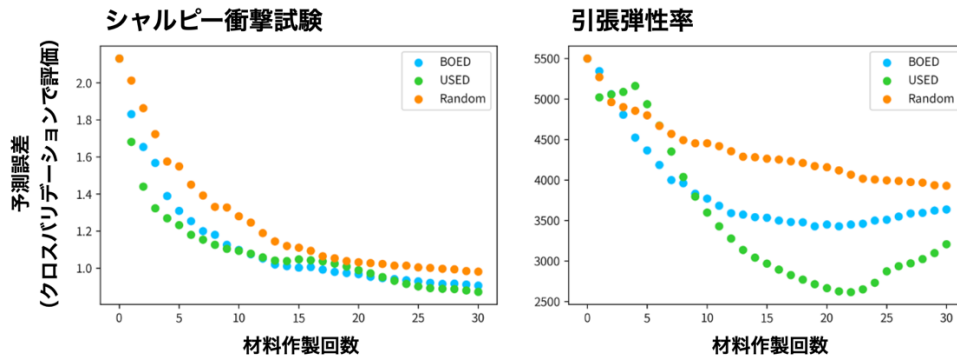


図3 シャルピー衝撃試験（左図）および引張弾性率（右図）に対する予測誤差の材料作製回数依存性。本研究で考案したベイズ最適化に基づく手法（BOED）および不確実性サンプリングに基づく手法（USED）、無作為に実験を行う手法（Random）を比較しています。Random に比べて、BOED および USED を用いることで少ない材料作製回数でも予測誤差が小さくなっていることがわかります。つまり、本研究で開発した新 AI 技術を利用することで、少ない作製回数で材料予測の精度向上を達成できました。

### 今後の展開

本研究では、XRD や DSC 等の測定データから材料物性の予測精度を高くする AI 技術を開発しました。この予測精度が向上すると、少ない材料作製回数で、材料の「構造」と「物性」の関係が明らかになり、物性の発現起源の明確化・材料開発指針の決定が可能となります。例えば、高分子材料の場合、高次構造と材料物性の関係を迅速に理解できるようになります。一般的に、高分子材料の機械物性は、一次構造<sup>3</sup>よりも高次構造に強く依存しており、この関係を理解することは、高分子材料開発において重要となります。

また、目的とする材料物性を測定する実験よりも、測定データが簡単に得られる場合、実験が簡単なデータから実験が困難なデータを予測することにつながり、材料設計の高速化と材料開発のコスト削減が実現できます。さらに、本手法は必要な実験回数を削減できるため、近年注目されている実験自動化技術と組み合わせることで、材料開発の高速化に貢献できると考えています。そのため、ここで開発した新 AI 技術は、材料開発の DX 基盤技術になると期待しています。

### 掲載論文

題目：Experimental design for the highly accurate prediction of material properties using descriptors obtained by measurement

著者：Ryo Tamura, Yuki Takei, Shinichiro Imai, Maki Nakahara, Satoshi Shibata, Takashi Nakanishi, and Masahiko Demura

雑誌：Science and Technology of Advanced Materials: Methods **1**, 152-161 (2021).

掲載日時：2021年9月28日（日本時間）オンライン掲載

### 用語解説

#### (1) ベイズ最適化 (Bayesian optimization)

機械学習を用いることで、対象とするスコアを大きくすると期待できる、次の実験を適切に選択する手法のこと。通常は、スコアとして材料物性そのものを採用し、目的の材料物性を実現するための AI 技術として利用されます。本研究では、このスコアを予測誤差に設定することで、機械学習の予測精度を良くするための手法として利用しています。

#### (2) 記述子

材料を表現するパラメータのこと。例えば、元素の種類を数値化したものや、材料を作成する際の温

度や圧力といったプロセスなども材料を表現するための記述子として利用されます。

(3) 一次構造と高次構造

高分子鎖は、分子が繰り返つながった構造をもち、これを特徴付ける構造を一次構造と呼びます。一方、高分子鎖が組織配列して形成される階層的構造を高次構造といいます。例えば、DNAにおいて核酸の塩基配列は一次構造であり、二重螺旋は高次構造に相当します。

(4) 不確実性サンプリング (uncertainty sampling)

機械学習の結果から、予測が最も不確かな場所を選択し、次に実験を行うことで機械学習の予測精度を良くする手法のこと。予測が不確かな場所は、一般的に実験データが少ない領域となります。

(5) クロスバリデーション

学習データを分割し、一部分を未知のデータとみなすことで、予測精度を評価する手法のこと。

### 本件に関するお問い合わせ先

国立研究開発法人物質・材料研究機構 経営企画部門 広報室

〒305-0047 茨城県つくば市千現 1-2-1

TEL: 029-859-2026, FAX: 029-859-2017

E-mail: [pressrelease@ml.nims.go.jp](mailto:pressrelease@ml.nims.go.jp)

旭化成株式会社 広報部 報道室

〒100-0006 東京都千代田区有楽町一丁目 1 番 2 号 日比谷三井タワー

TEL: 03-6699-3008

E-mail: [ak-houdou@om.asahi-kasei.co.jp](mailto:ak-houdou@om.asahi-kasei.co.jp)

三菱ケミカル株式会社 広報本部

〒100-8251 東京都千代田区丸の内 1-1-1

TEL: 03-6748-7161

E-mail: [MCHC-MBX-MCHC\\_HO\\_PR@mchcgr.com](mailto:MCHC-MBX-MCHC_HO_PR@mchcgr.com)

三井化学株式会社 コーポレートコミュニケーション部

〒105-7122 東京都港区東新橋一丁目 5 番 2 号 汐留シティセンター

TEL: 03-6253-2100

お問い合わせフォーム: [https://form.mitsuichemicals.com/corporate/cc\\_pr\\_csr\\_ja](https://form.mitsuichemicals.com/corporate/cc_pr_csr_ja)

住友化学株式会社 コーポレートコミュニケーション部

〒104-8260 東京都中央区新川 2-27-1 東京住友ツインビル (東館)

TEL: 03-5543-5102

E-mail: [sumika-kouhou@ya.sumitomo-chem.co.jp](mailto:sumika-kouhou@ya.sumitomo-chem.co.jp)