

Kernel SHAP を用いた高分子材料設計指針抽出の妥当性評価 Evaluating the Appropriateness of Extracted Polymeric Materials Design Guideline based on Kernel SHAP Explanation

Phua Yin Kan[†] 藤ヶ谷 剛彦[†] 加藤 幸一郎[†]
Yin Kan Phua Tsuyohiko Fujigaya Koichiro Kato

1. はじめに

近年、ビッグデータ、機械学習 (ML) 技術などの発達により、様々なデータの利活用で新たな価値を生み出し社会課題を解決する「データ中心社会」に進みつつある。同時に、UC Berkeley の研究者と Google DeepMind による無機材料が対象の A-Lab^[1]をはじめとした ML を活用したマテリアルズインフォマティクス (MI) の研究も盛んに行われている。^[2] エネルギー・環境・バイオなど社会生活を多方面で支える基幹材料群に属する機能性高分子への ML 導入も注目を集めているが、他材料と比べて ML の導入は立ち遅れている。^[3] 機能性高分子は一般的な高分子の主鎖骨格に、イオン伝導性など特定機能をもつ化学構造 (官能基) が側鎖として付加された材料である。構成する基本単位が 1 種であれば単独重合体、2 種以上であれば共重合体 (図 1) と呼ばれ、様々な構造体が存在する。そのため、実験過程は煩雑で危険な有機溶剤を多用しコスト・環境負担が重く、革新的で高効率な機能性高分子材料の開発を叶えるインフォマティクス技術の確立に期待がかかる。

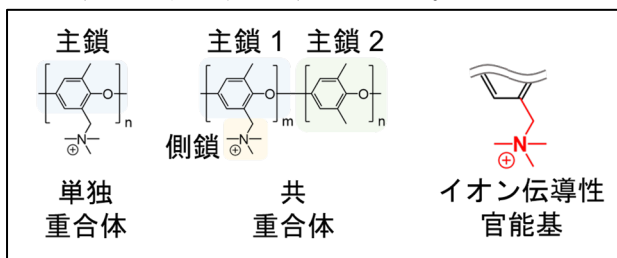


図 1 機能性高分子の構造

機能性高分子分野での導入が遅れている要因として、①文献などからの物性値の収集及び②単独・共重合体など異なる高分子構造を一元的に表現・管理可能なデータベース (DB) の構築が必要であり、参入障壁が高い。さらに③ ML の説明可能性が低くブラックボックス化していることが指摘されている。特に③については、ML から得られた結果の材料設計へのフィードバックを困難にしている。そこで、説明可能 AI (XAI) 手法の活用が材料分野においても注目を集めつつある。^[4] XAI 手法は ML モデルが予測の際に重視した説明変数を抽出することが可能な手法であり、抽出した説明変数を基に材料設計指針を導出し実験研究者と共有することで、材料開発の高効率化が期待される。

XAI の手法として、大きくモデル依存とモデル非依存に分けることが可能である。モデル依存手法として、random forest (RF) の feature importance や SHAP (SHapley Additive exPlanations) 内の TreeSHAP が知られており、モデル非依存手法では、Local Interpretable Model-Agnostic Explanations (LIME) や KernelSHAP が代表的である。Feature

importance は大局的、LIME は局所的な説明ができるが、SHAP は大局的・局所的な説明のいずれも可能である。機能性高分子材料分野では、少数データを基に ML を導入しているため、小規模なデータセットでも高精度な予測を可能とする決定木系モデルの RF を使った feature importance^[5]、または TreeSHAP を RF^[6]、eXtreme Gradient Boosting (XGBoost) ^[7,8] や Light Gradient Boosting Machine (LightGBM) ^[9] に導入した報告がなされてきた。最近ではデータ量の増加に伴い、深層学習の neural network (NN) を活用して KernelSHAP を導入した報告も出始めている。^[10] 将来的にデータ量が増加するにつれ、NN の利用、そして NN にも適用可能なモデル非依存の XAI 手法のひとつとして KernelSHAP の使用が拡大する可能性も高い。

しかし、XAI への関心が高まるにつれ、SHAP 値に基づいた説明が、説明変数と目的変数の本質的な関連性を必ずしも正しく反映していない可能性があるという懸念が報告された。^[11,12] Aas らの報告では、KernelSHAP が説明変数間の独立性を前提としている点に着目し、現実データにおいて多くの説明変数が相関関係にあることから、そのような仮定の下で算出された SHAP 値が非現実的かつ誤解を招く説明につながる危険性を報告した。^[11] また、Huang らによる報告では、KernelSHAP によって近似計算された SHAP 値のみならず、厳密に計算された正確な SHAP 値も誤解を招く可能性があることを示した。^[12] 特に、Huang らはデータセット内に目的変数と無関係な説明変数を挿入し、目的変数と相関する説明変数および無関係な説明変数の SHAP 値を計算したところ、実際に相関する説明変数の SHAP 値が 0 になり、無関係な説明変数は非ゼロになる事例を報告した。すなわち、SHAP から得られる重要度は本質的に誤った重要度である可能性があり、この問題は特に 4 種以上の説明変数を学習した分類器に顕著であると指摘した。

SHAP 値の算出により得られた各説明変数の重要度に基づいて材料設計指針を導出する手法は、これまで材料開発の効率化に資する有望なアプローチとして注目されてきた。仮に SHAP 値の信頼性への懸念が顕在化した場合、機能性高分子材料分野のみならず、材料分野全体における XAI の推進に対して深刻な課題を突きつける。MI によって抽出された重要説明変数が信頼に足る設計指針として活用することが困難になり、データ駆動型アプローチを通じた材料設計へのフィードバックが阻害される可能性がある。したがって、XAI 手法の信頼性評価は、MI の今後の発展において極めて重要な課題であると位置づけられる。

本研究では KernelSHAP を用いて抽出された高分子材料設計指針の妥当性を評価することを目的とする。SHAP により重要とされた説明変数が、適切に目的物性と関連し、化学的に意味のある因子であるかを検証する。この妥当性評価を通じて、今後 KernelSHAP の MI における設計指針抽出の信頼性担保に資する知見を得ることを目指す。

[†]九州大学 Kyushu University

2. 手法

本研究では機能性高分子材料の一種であるアニオン交換膜 (AEM) 高分子を対象に展開する。AEM 高分子は脱炭素社会の実現の鍵とされる燃料電池や水電解装置の中核部品である。しかし、AEM 高分子は現在性能指標であるアニオン伝導度や耐久性において課題があり、MI の導入かつ XAI 手法の適用による高アニオン伝導度・高耐久性の新 AEM 高分子の効率的な探索・発見に期待がかかる。アニオン伝導度とは、アニオン (負に帯電したイオン) が AEM 中を移動する能力を示す指標である。

2.1 データベースの構築

AEM 形燃料電池に用いられる AEM 高分子に関する公開 DB が未整備であるため、独自に AEM DB を構築した。^[8,13] データ収集にあたっては、2020 年前後に刊行された複数のレビュー論文を基礎とし、これらに掲載されている引用元論文を対象として情報を抽出した。データの抽出作業はすべて手作業で行い、各論文からはアニオン伝導度およびその測定温度、耐久性試験評価後のアニオン伝導度と対応する測定条件、AEM 高分子構造情報を文字・数値形式で収集した。また、論文中に掲載されている図表やグラフ中に存在する情報は、手動でデータ点を抽出した。AEM 高分子の化学構造については、分子構造を ASCII 文字列として表現する SMILES (Simplified Molecular Input Line Entry System) 記法を用いて文字列データとして抽出した。単独・共重合体 (図 1) が存在する AEM 高分子の化学構造情報は、繰り返しユニットが 2 種類存在する共重合体をベースに、それぞれの繰り返しユニットをブロック A と B に分け、1 種類の繰り返しユニットのみ有する単重合体はブロック A と B それぞれに同じ化学構造を入力することで一元的に収録した (図 2)。^[8,13]

	ブロック A	ブロック B
単重合体		
モル比	0.5	0.5
共重合体		
モル比	0.758	0.242

図 2 AEM 高分子構造の繰り返しユニットの収録方法

また、AEM 高分子のイオン交換容量 (IEC) は、アニオン伝導度に影響を与える重要な構造指標の一つであるため、以下の式で算出してデータベースに加えた。

$$IEC = 1000 \times \frac{\text{カチオンのモル数}}{\text{繰り返しユニットの平均分子量}} [\text{mequiv. g}^{-1}]$$

カチオンのモル数はプラスの電荷をもつイオンの量であり、繰り返しユニットの平均分子量は「ブロック A の分子量」

×「ブロック A モル比」+「ブロック B の分子量」×「ブロック B モル比」である。

2.2 データの前処理

化学構造情報を SMILES から ML モデルが学習しやすい数値形式に変換するため、2018 年までに存在する記述子 (化学構造を数値化したもの) への変換が可能な Python の公開ライブラリである Mordred^[14]を用いて、DB 内に含まれる全 AEM の化学構造を記述子変換した。得られた記述子から非数値列や全ての列で 0 を示す記述子を削除した。

説明変数間の相関は Pearson Correlation Coefficient にて算出し、0.1~1.0 の間で 0.1 刻みに閾値を設定して説明変数を選別することで 10 種の DB を構築した。具体的には、説明変数 A と説明変数 B との相関係数が閾値以下の場合には両者を残し、閾値を超えた場合は A と B のそれぞれについて全説明変数との相関係数を算出し、平均相関係数が低い方を残すという操作を全説明変数ペアに対して行った。さらに記述子の半数が 0~1.5 の範囲内の数値をもっている特性に合わせて、SHAP 値の信頼性にストレスがかかりうる設定として 0~1.5 のレンジで無関係の数値を生成しそれらを含む列 (random feature) を各 DB に追加した。得られた DB を閾値に応じて「DB_{i}」(i=0.1, 0.2, ..., 1.0) と呼ぶ。同様の処理を別途作成したテスト DB にも適用した。

2.3 ML モデルの構築

ML モデルには高速性と予測精度が高く、TreeSHAP および KernelSHAP の双方が適用可能な XGBoost 回帰モデルを使用した。XGBoost のハイパーパラメータチューニングは Optuna で行なった。交差検証 (CV) の設定は group k-fold (k=5) であり、同一高分子構造からは同一 IEC が得られるため、group は IEC に基づいて分割した。これは、同一構造で実験条件違いのデータが学習・検証双方に意図せず漏洩することを防ぐためである。学習検証精度は決定係数 (R²) で評価した。目的変数はアニオン伝導度とし、DB_{i} 毎にモデルを構築した。

2.4 予測論理評価

構築した XGBoost の予測論理及び SHAP の信頼性を評価するため、各 DB_{i} で学習検証して得た XGBoost_{i} に対して、TreeSHAP 及び KernelSHAP を適用した。信頼性問題のひとつである無関係の説明変数への非 0 の SHAP 値算出を検証するため、TreeSHAP・KernelSHAP から得られた大局的な説明変数の重要度上位にて無関係の説明変数が入り込んでいるかを解析した。

3. 結果と考察

3.1 データベースの構築

本研究では独自に構築済の 78 報の論文から計 2,452 点のアニオン伝導度関連データを抽出した DB を用いた。

3.2 データの前処理

Mordred 記述子ライブラリを用いることで、各 AEM 高分子構造に対して 3,226 次元の説明変数を得た後、非数値列及び全 0 列を削除した結果、説明変数は 2,231 次元となった。これらに実験条件を加えることで、説明変数は

2,240次元となった。説明変数間の相関に応じた説明変数の選別の結果、閾値が0.1~1.0(絶対値)の場合、説明変数の次元はそれぞれ9、15、26、35、54、96、169、304、567、2,228となった。各閾値で選別した説明変数に random feature を足し、目的変数と紐づけることで $DB_{\{i\}}$ を得た。

3.3 ML モデルの構築

XGBoost 回帰モデルを構築した結果、Pearson 相関係数の閾値が0.7の時に $R^2 = 0.73$ と最も高いことが分かった(図3)。 $DB_{0.7}$ を学習検証する際に Optuna で最適化されたハイパーパラメータは $n_estimators = 640$ 、 $max_depth = 8$ 、 $learning_rate = 0.136422504$ 、 $min_child_weight = 14$ 、 $subsample = 0.441589551$ 、 $colsample_bytree = 0.772505702$ 、 $reg_alpha = 0.010344708$ 、 $reg_lambda = 1.206544598$ 、 $gamma = 0.000840019$ であった。全体として、 $DB_{\{i\}}$ の i が増加するにつれ、 R^2 が高くなる傾向があり、学習検証精度は $i = 0.3$ で $R^2 = 0.93$ という高いレンジに入り、テスト精度は $i = 0.7$ でピークアウトすることが判明した。

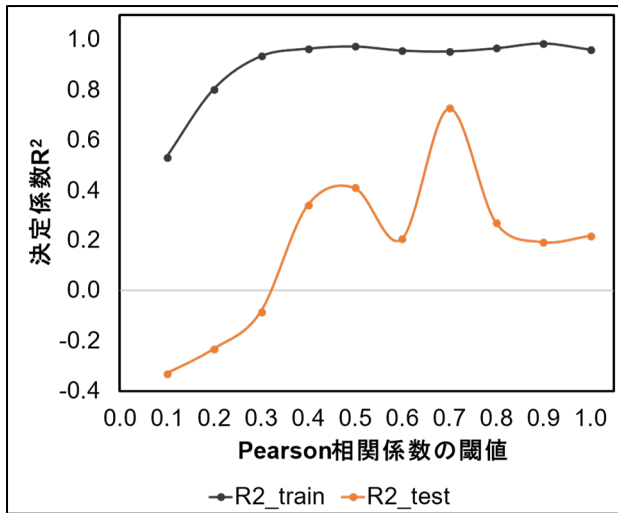


図3 $DB_{\{i\}}$ を基に構築した XGBoost $\{i\}$ の学習検証(灰色)・テスト(橙色)予測精度

3.4 予測論理評価

SHAP 値の信頼性評価では、TreeSHAP は近似ではなく正確な SHAP 値算出手法が用いられているため、^[15] TreeSHAP をコントロールとし、KernelSHAP を評価対象とする。全 $DB_{\{i\}}$ に IEC は説明変数として含まれており、実験的にも真にアニオン伝導度と相関があることが既知であるため、random feature との比較対象として IEC を用いることとした。各閾値 i での IEC 及び random feature の大局的な順位を抽出したところ、相関係数の閾値 i を低く設定している場合、特に $i = 0.1$ の際は、残存説明変数の数が少ないため、random feature にも高い SHAP 値が算出され、順位も TreeSHAP・KernelSHAP 問わず 5 位に位置した(図4a)。一方、IEC の順位は TreeSHAP・KernelSHAP を問わず 1 位であったため、この結果からは SHAP 値の信頼性の高低は結論づけられない。閾値 i が 0.2 から 0.6 にかけて、IEC は TreeSHAP・KernelSHAP 問わず 1・2 位を維持しているのに対して、random feature は TreeSHAP の場合、順位はそれぞれ $7 \rightarrow 16 \rightarrow 22 \rightarrow 39 \rightarrow 78$ と、急降下した。一方、

KernelSHAP から算出した順位も $7 \rightarrow 16 \rightarrow 22 \rightarrow 38 \rightarrow 94$ と急降下しているが、特に $i = 0.6$ では、TreeSHAP より random feature の順位が下がった(図4a)。IEC の順位は TreeSHAP・KernelSHAP 問わず $i = 0.9$ まで 2 位を維持したままであり、唯一 KernelSHAP は $i = 1.0$ の際に IEC の順位が 3 位まで下落した。とはいえ、誤差範囲内と捉えられるため、IEC の順位差からは明確に TreeSHAP・KernelSHAP の違いは確認できなかった。Random feature の場合、TreeSHAP が算出した順位はそれぞれの i で $118 \rightarrow 121 \rightarrow 33 \rightarrow 121$ であり、KernelSHAP の場合はそれぞれ $119 \rightarrow 141 \rightarrow 38 \rightarrow 141$ である。これらの順位からは、KernelSHAP は TreeSHAP よりも無関係な説明変数の順位を引き上げない傾向があることが分かった。

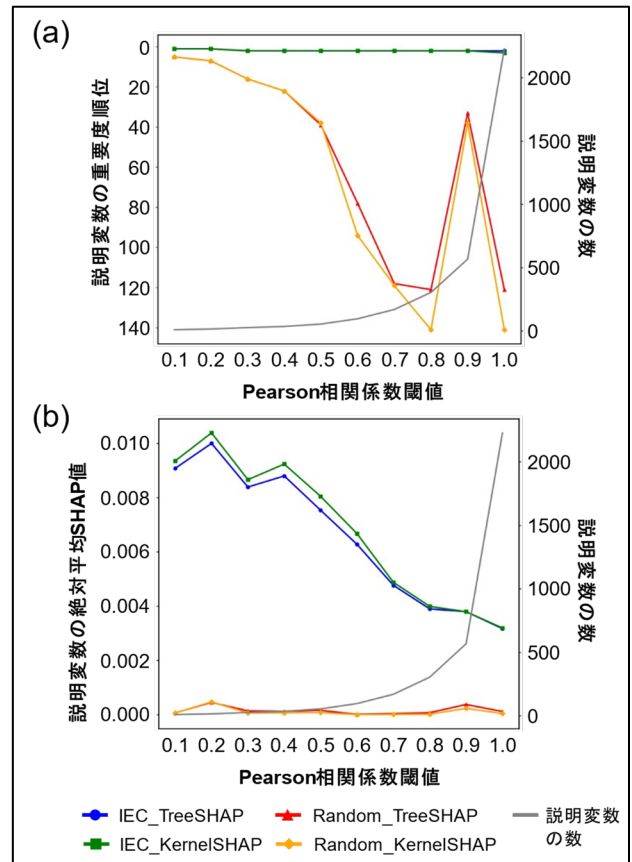


図4 Pearson 相関係数閾値 i ごとの (a) 説明変数の重要度順位; (b) 説明変数の絶対平均 SHAP 値

IEC・random feature の絶対平均 SHAP 値を解析したところ、大局的な順位とおおむね同様な傾向が得られた(図4b)。 i が 0.7 に近づくにつれ、IEC における TreeSHAP・KernelSHAP の絶対平均 SHAP 値の差が縮まり、random feature では IEC ほどの差はないが、同様な傾向が見られた。 $i > 0.7$ の場合、特に $i = 1.0$ の差が顕著であり、TreeSHAP における random feature の絶対平均 SHAP 値=0.000115 に対し、KernelSHAP からは 3.86×10^{-5} と、2桁の差がついた。SHAP 値の信頼性を議論する際の基準は、無関係な説明変数には 0 の SHAP 値を与えているか否かを設定している^[12] ため、今回の結果からは 0 に近い SHAP 値を与えた KernelSHAP の方が信頼できることが示唆された。

さらに、XGBoost の学習検証精度による影響を考慮するため、学習検証精度もテスト精度も高く、最もバランスよく学習している $i = 0.7$ で比較した結果、若干 KernelSHAP の方が 0 に近く信頼性が高い結果となった (表 1) が、無関係な説明変数が上位に入る恐れはないと考えられる。

表 1 $i = 0.7$ 時の説明変数別 SHAP 順位・値

説明変数_SHAP 種類	順位 /169	絶対平均 SHAP 値
IEC_TreeSHAP	2	0.00477
IEC_KernelSHAP	2	0.00488
random_TreeSHAP	118	3.72×10^{-5}
random_KernelSHAP	119	5.11×10^{-6}

現在の MI 分野における研究動向を踏まえ、説明変数の数が 20 を超える場合であっても、実際に解釈の対象とされるのは上位 20 変数程度に限定されていることが多い。本研究結果からは $i = 0.4$ 以降で説明変数の数が 20 以上である場合、random feature 上位 20 位以内に入ることは確認されておらず、現在の MI 手法に対する実質的な影響は限定的である。また、計算コスト問題の考慮及び決定木系モデルを使用した場合、計算コストが低く幅広い計算機器に対応できる TreeSHAP ではなく KernelSHAP を使用する必要性が低い。しかし、今後の MI 手法の高度化によって上位 20 変数に限らず、より広範な説明変数も解釈対象に含まれる可能性がある。その際、モデル非依存かつ柔軟な KernelSHAP が優位になることも想定される。

SHAP で得た上位 20 に含まれる説明変数は、実験条件など既知の説明変数以外は記述子由来の説明変数である (図 5)。これらの記述子は、既報にて説明可能であることが示されており、^[8,13] 材料設計指針への導出も見込まれる。TreeSHAP・KernelSHAP 問わず、20 以上の説明変数がある場合、無関係な説明変数が上位に入り込む可能性が低く、導出した材料設計指針の信頼性も担保できる。

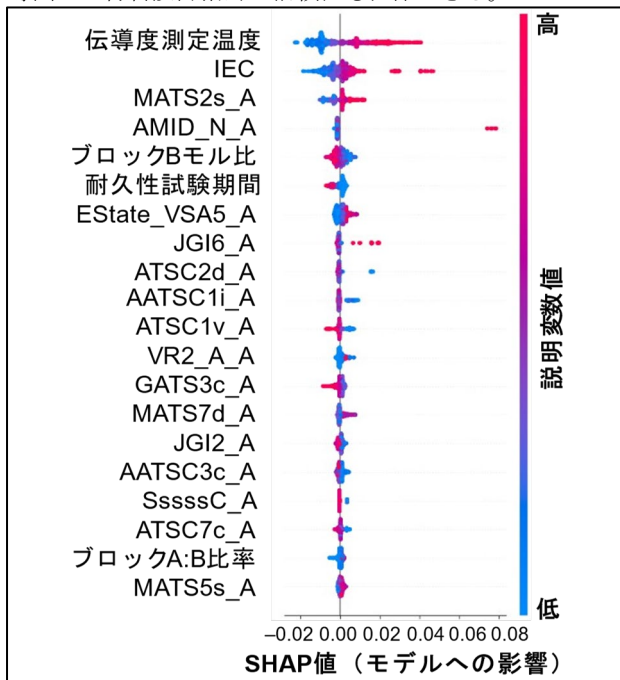


図 5 $i = 0.7$ の TreeSHAP summary plot

4. おわりに

本研究では TreeSHAP をコントロールに、KernelSHAP の信頼性評価を行なった。重要かつ目的変数と関連していることが既知である IEC 及び無関係の random feature への SHAP 値を比較・解析した結果、KernelSHAP は TreeSHAP より random feature の SHAP 値を抑える性能に優れていることが判明した。これは、今後機能性高分子のみならず、分子記述子を活用している材料分野の MI が NN 等複雑なモデルに移行した際、XAI 手法として KernelSHAP の有用性を示唆する結果である。

謝辞

本研究は、JST ACT-X (JPMJAX22AF)、JST 科学技術イノベーション創出に向けた大学フェローシップ創設事業 (JPMJFS2132) と JST 次世代研究者挑戦的研究プログラム (JPMJSP2136) の支援を受けたものです。

参考文献

- [1] N. J. Szymanski, B. Rendy, Y. Fei, R. E. Kumar, T. He, D. Milsted, et al., "An autonomous laboratory for the accelerated synthesis of novel materials", *Nature*, Vol. 624, No.7990 (2023).
- [2] G. J. Mulholland, S. P. Paradiso, "Perspective: Materials informatics across the product lifecycle: Selection, manufacturing, and certification", *APL Materials*, Vol.4, No.5 (2016).
- [3] W. Ge, R. De Silva, Y. Fan, S. A. Sisson, M. H. Stenzel, "Machine Learning in Polymer Research", *Advanced Materials*, Vol. 37, No. 11 (2025).
- [4] X. Zhong, B. Gallagher, S. Liu, B. Kailkhura, A. Hiszpanski, T. Y.-J. Han, "Explainable machine learning in materials science", *npj Computational Materials*, Vol. 8, No. 1 (2022).
- [5] E. J. Di Mare, A. Punia, M. S. Lamm, T. A. Rhodes, A. J. Gormley, "Data-Driven Design of Novel Polymer Excipients for Pharmaceutical Amorphous Solid Dispersions", *Bioconjugate Chemistry*, Vol. 35, No. 9 (2024).
- [6] H. Okuyama, Y. Sugawara, T. Yamaguchi, "Machine-Learning-Aided Understanding of Protein Adsorption on Zwitterionic Polymer Brushes", *ACS Applied Materials & Interfaces*, Vol. 16, No. 19 (2024).
- [7] B. Esteki, M. Masoomi, M. Moosazadeh, C. Yoo, "Data-Driven Prediction of Janus/Core-Shell Morphology in Polymer Particles: A Machine-Learning Approach", *Langmuir*, Vol. 39, No. 14 (2023).
- [8] Y. K. Phua, T. Fujigaya, K. Kato, "Predicting the anion conductivities and alkaline stabilities of anion conducting membrane polymeric materials: development of explainable machine learning models", *Science and Technology of Advanced Materials*, Vol. 24, No. 1 (2023).
- [9] M. Yang, J. J. Zhu, A. L. McGaughey, R. D. Priestley, E. M. V. Hoek, D. Jassby, et al., "Machine Learning for Polymer Design to Enhance Pervaporation-Based Organic Recovery", *Environmental Science & Technology*, Vol. 58, No. 23 (2024).
- [10] S. Glass, M. Schmidt, P. Merten, A. Abdul Latif, K. Fischer, A. Schulze, et al., "Design of Modified Polymer Membranes Using Machine Learning", *ACS Applied Materials & Interfaces*, Vol. 16, No. 16 (2024).
- [11] K. Aas, M. Jullum, A. Løland, "Explaining individual predictions when features are dependent: More accurate approximations to Shapley values", *Artificial Intelligence*, Vol. 298 (2021).
- [12] X. Huang, J. Marques-Silva, "On the failings of Shapley values for explainability", *International Journal of Approximate Reasoning*, Vol. 171 (2024).
- [13] Y. K. Phua, N. Terasoba, M. Tanaka, T. Fujigaya, K. Kato, "Unsupervised Machine Learning - Derived Anion - Exchange Membrane Polymers Map: A Guideline for Polymers Exploration and Design", Vol. 11, No. 14 (2024).
- [14] H. Moriwaki, Y. S. Tian, N. Kawashita, T. Takagi, "Mordred: a molecular descriptor calculator", *Vol. 10, No. 1 (2018)*.
- [15] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, et al., "From Local Explanations to Global Understanding with Explainable AI for Trees", Vol. 2, No. 1 (2020).