

機能性高分子の物性予測に適用可能な説明可能ニューラルネットワークの構築

Establishing Explainable Neural Network for Predicting Physical Properties of Functional Polymer

九大工¹, 九大 CMS², I²CNER³, 九大 RIIT⁴ °Phua Yin Kan¹, 藤ヶ谷 剛彦^{1,2,3}, 加藤 幸一郎^{1,2,4}
Kyushu Univ. Grad. Sch. of Eng.¹, Kyushu Univ. CMS.², I2CNER³, Kyushu Univ. RIIT⁴, °Yin Kan
Phua¹, Tsuyohiko Fujigaya^{1,2,3}, Koichiro Kato^{1,2,4}

E-mail: fujigaya.tsuyohiko.948@m.kyushu-u.ac.jp, kato.koichiro.957@m.kyushu-u.ac.jp

近年材料インフォマティクス(MI)の発展と共にニューラルネットワーク(NN)の高予測精度が注目され、MI 導入で NN を使用する事例が増えている[1]。しかし、NN のブラックボックス問題から多くの事例で NN は物性予測の活用に留まっており、構築した NN から新しい材料設計指針が得られず MI の利点を生かし切れていない。実際、機能性高分子材料の一種であるアニオン交換膜(AEM)においても、NN 導入事例が複数ある一方、NN のブラックボックス問題に言及していない。我々は、AEM 材料を対象に決定木系モデルを基に SHAP 値[2]を活用した説明可能機械学習(ML)モデルの構築し、SHAP 値解析で抽出した重要説明変数は新材料設計指針になりうることを示した[3]。本研究では、AEM 材料を対象に NN を構築し、各説明変数の SHAP 値を基に重要説明変数を特定し新材料設計指針を抽出できる説明可能な NN ワークフローの確立を目指す。

本研究では 78 報分の論文データからなるデータベース(DB)を独自で構築して使用した[3]。AEM 化学構造の記述子変換には Mordred library [4]を用いた。NN の学習には、目的変数をアニオン伝導度、説明変数は記述子化した化学構造及び実験条件を用いた。NN で SHAP を算

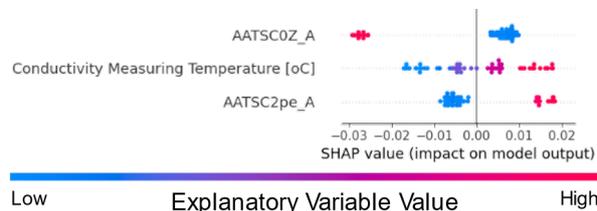


図 1. NN がテストデータを予測する際に重要視した説明変数の SHAP 図.

出する報告例がないのはメモリコストが高いことが考えられ、本研究では高次元の記述子に対し次元削減することでメモリコストを抑えた。学習検証精度は決定係数 R^2 で評価し、0.814 と高い精度を得た。構築済 NN で検証データと異なるテストデータを予測させた精度は 0.602 と、実用上問題ない精度を得た。テストデータを予測する際に重要視した説明変数を SHAP 値解析した結果、記述子由来の AATSC0Z_A が低い値をもつことが重要であることが判明した(図 1)。AATSC0Z は元素種類の多様性を評価する記述子であり、分子構造内に多種類・高い原子量元素をもつことが AATSC0Z_A 値に高く寄与する。AEM 材料内における高い原子量元素は主にハロゲン系であり、ハロゲン系元素の除去によるアニオン伝導度向上という新材料設計指針になりうる知見を得た。

[1] C. Li et al., *InfoMat*, **2023**, 5, 8, e12425. [2] S.M. Lundberg et al., *Proceedings of the 31st International Conference on Neural Information Processing Systems*, **2017**, 4768-4777. [3] Y.K. Phua et al., *Sci. Technol. Adv. Mater.*, **2023**, 24, 1, 2261833. [4] H. Moriwaki et al., *J. Cheminformatics*, **2018**, 10, 4.