

信頼される AI の基盤技術  
2021 年度採択研究代表者

2022 年度  
年次報告書

横川 大輔

東京大学 大学院総合文化研究科  
准教授

化学的知見を生かした転送性の高い特徴量の抽出と利用

## 研究成果の概要

本研究課題は「化学的知見を生かした転送性の高い特徴量の抽出と利用」であり、得られた結果に対する説明性の解明に重きを置いている。2022年度は、①これまでに構築した学習モデルの畳み込み&連結層の再検討、②水一オクタノール分配係数(logP)に対する学習とその説明性に関する検討を行った。以下で詳細に説明する。

### ① これまでに構築した学習モデルの畳み込み&連結層の再検討

本研究では、グラフニューラルネットワーク(GNN)に基づいたモデルを用いて検討を行っている。GNNを用いた学習モデルの適用は化学においても盛んに行われており、特にラプラシアンを用いた畳み込みを用いたモデルは様々な対象に対し適用されてきた。この方法は多くの場合において良い予測性能を示す一方で、畳み込みの層数をある一定以上増やしても精度の向上が見られない問題点が指摘されてきた。本研究ではこの畳み込みに注目し、分子の結合数ごとに情報を連結する新たな畳み込み法を提案した。この新規手法を従来のラプラシアンを用いた畳み込みと比較したところ、 $N$ 層目の畳み込みで追加される情報と $N-1$ 層目までに追加された情報との間の冗長性が大幅に削減されていることが明らかになった。このため、新規畳み込みを用いた本モデルの場合、畳み込みの層数を増やしても予測性能は低下することがないことを明らかにした。

### ② 水一オクタノール分配係数(logP)に対する学習とその説明性に関する検討

①で構築したモデルを用い、logPの予測を行った。本手法の精度をこれまでに報告されているlogPの予測プログラムの結果と比較すると、ほぼ同程度の結果が得られることがわかった。この構築した学習モデルに対し、Integrated Gradientsを用いた解析を行ったところ、logPの予測が化学で用いられる置換基と似た情報を経て行われていることが明らかとなった。