

分子設計における AI 活用の動向と将来展望

アステラス製薬株式会社

藤 秀義

近年、深層学習をはじめとする人工知能（AI）技術の急速な発展に伴い、自動運転や医療画像診断など様々な分野での実用化が進んでいる。医薬品の分子設計の分野においても深層学習を応用した事例が多く報告されている。2012 年 Kaggle で開催された Merck Molecular Activity Challenge¹での活性予測、2014 年に NIH 主催で開催された Tox21 Data Challenge 2014²での毒性予測、いずれのコンテストにおいても深層学習を用いたチームがトップの成績を収めている。最近では、化合物の構造発生、逆合成解析における深層学習の活用も進められており、どんな化合物を作るのか、どうやって作るのかまで、分子設計の一連のプロセスを AI で予測可能になってきている。

深層学習の優れた点は、人間の手を借りずとも、学習データの中から特徴量を自動的に抽出可能な点である。それは逆に言えば、どんな特徴量を抽出するかは AI 任せであり、AI が何を根拠に判断しているかが分からないというのが問題である。これはブラックボックス問題と呼ばれており、深層学習の懸念点のうちの 1 つである。このブラックボックス問題を解消しようと、説明可能な AI（Explainable AI: XAI）の研究が既に進められている。分子設計においても、何を作るのか、どうやって作るのかに加え、なぜそれを作るのかの判断根拠を求められることがあり、説明可能な AI 開発は興味深い研究課題である。

化学構造式は化学者にとっての共通言語であり、化学構造式 1 つには反応性や毒性など多くの情報が含まれている。メディシナルケミストは過去の経験・知識を基に、化学構造式からそれらの情報を読み解き、分子設計を進めることが可能である。それではもし AI が化学構造式から、その分子の物性や毒性を予測出来るようになった場合、AI は判断根拠を示すことができるのだろうか。本発表では、説明可能な AI 技術を毒性予測に応用し、化学的に意味のある根拠を示すことができるのかを検証した結果を紹介したい。

1. <https://www.kaggle.com/c/MerckActivity>
2. <https://tripod.nih.gov/tox21/challenge/>