

2025年8月8日

報道関係者各位

北里大学

世界最大級の化合物データに対応可能な 革新的な分子生成 AI モデルを開発 ～創薬研究・化学物質開発の飛躍的加速に期待～

北里大学未来工学部データサイエンス学科の榊原康文教授の研究グループは、膨大な化合物データを用いて高精度な分子構造生成を可能にする、世界最大規模の分子生成 AI モデル「FRATTVAE (Fragment Tree-Transformer VAE)」を開発しました。この成果により、従来の生成モデルが扱えなかった数千万規模の分子データセットの学習と、高度に複雑な分子構造の効率的生成が可能となります。本研究成果は2025年8月5日付で、化学分野の国際的ジャーナル『Communications Chemistry』に掲載されました。

研究成果のポイント

- ◆世界最大規模となる1200万種類以上の分子データを学習可能なAIモデルを開発
- ◆分子の部分構造(フラグメント^{【*1】})をツリー型データとしてTransformer^{【*2】}で処理することで計算効率を劇的に向上
- ◆従来モデルを大幅に上回る精度・効率で新規かつ実用的な化学物質を生成可能

研究の背景

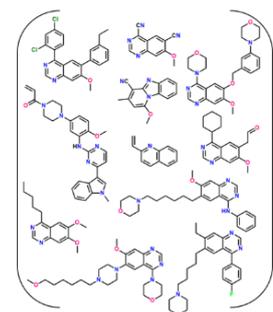
医薬品開発や材料科学の分野において、望ましい性質を持つ新規化合物の探索には莫大な時間とコストがかかります。近年では、化合物をAIで自動的に生成する「分子生成モデル」が注目されています。しかし、従来のAIモデルはデータサイズや分子構造が大きくなると精度が低下したり、計算が困難になる問題を抱えていました。そこで本研究グループは、大規模で複雑な分子データに効率よく対応できる新たなAIモデルの開発に取り組みました。

研究内容と成果

本研究で開発したFRATTVAEは、分子をフラグメント(部分構造)の集合として扱い、それらをツリー構造で表現することで分子の複雑さに対応します。また、自然言語処理の最新AI技術「Transformer」を用い、フラグメント間の複雑な関係を自己注意メカニズム(Self-Attention)でモデル化します。これにより並列計算が可能になり、従来モデルの制約であった大規模データセットの処理を実現しました。本研究ではChEMBL、DrugBank、PubChemなどのデータベースから収集した1200万種類を超える化合物を用いてFRATTVAEを学習しました。その結果、FRATTVAEは既存手法と比較して分子構造の再現精度、生成された分子の妥当性および新規性などの各種指標で非常に高い性能を示しました。また、医薬品開発に関連する重要な性質(薬らしさ、合成容易性、天然物らしさ)においても、既存手法を一桁以上上回る性能*を達成しました。

* Reconstruction(再構成精度)、NP(天然物らしさ)、QED(経口薬らしさ)の数値において

・ AI創薬
生成AIによるバーチャル
化合物構造の生成



今後の展開

FRATTVAE は創薬分野をはじめ、幅広い化学・材料分野での応用が期待されます。特に、従来技術では扱えなかった大規模かつ複雑な分子データを基にした創薬研究や新素材開発を加速させることが可能となります。今後は、さらに大規模な化学データベースを用いてモデルを発展させ、産学連携を通じた実用化研究を進める予定です。

論文情報

掲載誌：Communications Chemistry

論文名：Leveraging Tree-Transformer VAE with fragment tokenization for high-performance large chemical model generation

著者：Tensei Inukai, Aoi Yamato, Manato Akiyama, Yasubumi Sakakibara

DOI： [10.1038/s42004-025-01640-w](https://doi.org/10.1038/s42004-025-01640-w)

■本研究は文科省科研費学術変革領域研究（A）「天然物が織り成す化合物潜在空間が拓く生物活性分子デザイン」23H04885, 23H04880 の助成を受けたものです。

用語解説

※1 フラグメント (Fragment)

化学分野では、分子の一部を構成する特定の小さな構造を指す。

※2 Transformer (トランスフォーマー)

自然言語処理で開発された深層学習モデル。自己注意 (Self-Attention) 機構により並列処理が可能となり、効率的に大規模データを扱えるのが特長である。

問い合わせ先

《研究に関すること》

北里大学 未来工学部 データサイエンス学科 人工知能研究室

教授 榊原 康文 (さかきばら やすぶみ)

e-mail： sakakibara.yasubumi@kitasato-u.ac.jp

URL： <https://www.kitasato-u.ac.jp/fr-eng/laboratory/ai/>

《取材に関すること》

学校法人北里研究所 広報室

〒108-8641 東京都港区白金 5-9-1

TEL：03-5791-6422

e-mail： kohoh@kitasato-u.ac.jp