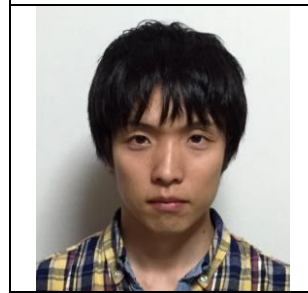


データサイエンス

key word

- メタボローム
- 機械学習
- 薬物動態
- データキュレーション



江崎 剛史
Tsuyoshi Esaki

データサイエンス学部
准教授

- 【プロフィール】**
- ・2008年 岡山大学 理学部数学科 卒業
 - ・2010年 岡山大学大学院 自然科学研究科 修了
 - ・2016年 大阪大学大学院 生命機能研究科 修了 博士(理学)
 - ・2010年 - 2012年 広島大学大学院医歯薬学総合研究科運営支援グループグループ員
 - ・2012年 - 2016年 理化学研究所生命システム研究センター テクニカルスタッフ
 - ・2016年 - 2019年 医薬基盤・健康・栄養研究所 特任研究員
 - ・2019年 - 2020年 滋賀大学データサイエンス教育研究センター 助教
 - ・2020年 - 2023年 同 准教授 (2023年～データサイエンス学部 准教授)

- 【主な社会的活動】**
- ・情報計算化学生物(CBI)学会
 - ・日本薬物動態学会
 - ・構造活性相関部会
 - ・CBI学会年会プログラム委員

【主な論文】

T. Esaki, et al., Data Curation can Improve the Prediction Accuracy of Metabolic Intrinsic Clearance, *Molecular Informatics*, 2019, e1800086

【代表的な研究テーマ】
□ 創薬の効率化を目指した情報科学手法の活用

課題解決に役立つシーズの説明

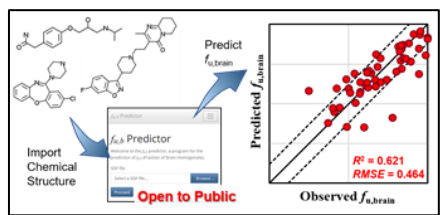
創薬には膨大な労力と費用が掛かっており、1つの医薬品の研究開発にかかる費用は2500億円以上、15年以上もの歳月がかかるとの報告がある。従来から行われている手順として、まず標的分子に作用する候補化合物を購入・合成し、吸収や代謝、副作用等を評価する試験を行い、効果が低い、もしくは副作用が生じる可能性のある化合物を除外する。そして残った化合物の構造を改善して試験で評価を行うという作業を幾度となく繰り返し、新薬として創り上げていく。創薬の初期段階では、薬の可能性のある化合物を可能な限り用いて試験を行うため、試験手法が多岐にわたり、非常に多くの労力と費用が費やされる。

近年は創薬の効率化を目指し、化合物の化学構造だけから薬としての特性を予測する機械学習モデルが注目されている。化合物の構造だけから特性を予測できれば、新薬となる可能性の高い化合物を優先的に試験でき、労力や費用の大幅な削減が期待できるためである。しかし、より効率的に創薬を進めるためには、予測だけでなく「薬になりにくい化合物をどのように変換したら薬になり得るか」を考える手がかりを得ることも重要である。

そこで現在は、「機械学習による予測モデル構築」と「予測結果の解釈に向けた特徴量選択」、「データ収集のシステム化」の3点からのアプローチにより、創薬の効率化を目指した研究を行っている。

■ 薬物動態に関する性質を予測する機械学習モデルの構築

創薬の早期段階で化合物の特性を予測することは、効率的な創薬に非常に重要である。情報技術の進歩により、様々な薬物動態パラメータを予測するモデルが開発されてきたが、ほとんどのパラメータは商用ソフトウェアでのみ使用可能である。そこで、化合物の構造情報のみから薬物動態を予測できる機械学習モデルの開発を行っている (Esaki et al. 2019, *Journal of Chemical Information and Modeling*, accepted, 右図)。

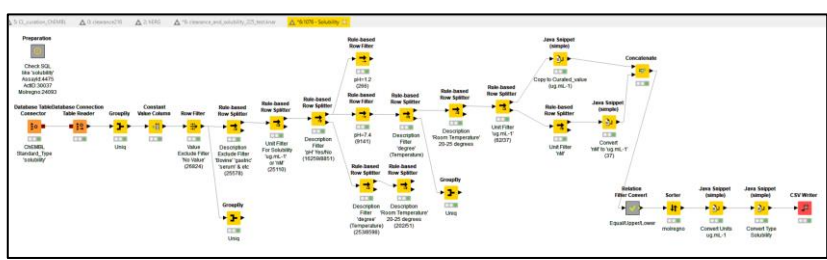


■ 医薬品候補の構造改善を提案する手法の開発

複雑な試験の結果を高い精度で予測するためには、大量の特徴量を使って非線形の機械学習モデルが構築されることが一般的であり、モデルの解釈が非常に困難となる。そこで「非線形スパース正則化を用いた医薬品候補の構造改善を提案する手法の開発」とし、解釈可能な範囲に特徴量を削減すると同時に、予測モデルの精度を高く維持する手法の検討を行っている。

■ 公共データの収集・前処理を簡略化するシステムの開発

予測モデルの構築には大量のデータが不可欠であるが、研究室単位で取得できるデータの数には限りがあるため、大規模データの収集元として、公共データベースが広く活用されている。しかし、公共データベースには転記ミスや情報の不足が多く見受けられ、データを精査して収集することが必要となる。そこで、公共データベースから求めるデータを抽出し、解析に使える形に前処理を行う自動システムの開発(下図)を進めている。



企業・自治体へのメッセージ

主に創薬に関するデータを扱ってきたため、創薬や医療に近い領域からの共同研究を希望しますが、使用してきた手法は創薬だけにとどまらず、広い分野においても適用できると考えています。