

AI 創薬に立ちはだかる壁とその超え方 How to overcome the barriers to AI drug discovery

三澤 大太郎
Daitaro Misawa

SyntheticGestalt 株式会社
SyntheticGestalt

本講演では、Lead Generation に機械学習を活用する際の課題について取り上げ、その対策について述べる。具体的には、評価時と利用時のモデルの振舞いが異なることにより、予測結果通りに実際のアッセイ結果が得られない問題について考察する。

機械学習は本質的に普遍性を持ち得ないゆえに理論的動作保証が難しく、せいぜいモデルの汎化性能を向上させることが実用上のプラクティスである。更に、汎化性能評価は評価設計に依存しており、ナイーブな方法では間違った意思決定を導く危険性がある。例えば、機械学習の標準的プロトコルを単純に用いた結果、実際以上に高いパフォーマンスを観測してしまい、デプロイすべきモデルを誤って選択することなどが挙げられる。

この問題を回避するためには、創薬という事業ドメインを考慮して評価手法をカスタマイズする必要がある。特に、機械学習ベースのスクリーニングにおける推論データは、各種化合物ライブラリなどで構成されることが多く、推論データの分布は学習データの分布とは一般に異なるため、テスト評価はリハーサルとなるように調整しなくてはならない。従って、重要な施策の一つは学習時と推論時のデータセットの素性を似せることである。しかし、実際のリソースの入手状況下では学習データは十分に豊かでないため、現実的には、テストで得られるモデル適用可能領域 **Applicable Area** に合わせて推論データを構成することになる。

モデル適応可能領域を見定め、かつ、リハーサルとして現実の一部環境を模すため、少なくとも下記 2 つの評価手法を導入することが望ましいと考えている。1 点目は、評価データを様々な商集合構成によって細かくセグメント化し、各ブロック上でパフォーマンス評価を行ってモデルがワークする部分集合の条件を見定めていく **Partitioned Evaluation** である。2 点目は、確率モデルが必ず内包している分散の影響を見積もるための簡易的手法であり、現実的状况と同様に過負荷な **Rare Positive** の設定を作り出す **Stressed Evaluation** である。