日時: 2021年10月25日13:00-17:00

半日で知る、化学分野のデータサイエンスおよび人工知能概要: 「FS-08:化学データサイエンスおよび人工知能討論、勉強会」立ち上げ会協賛 A Half-day Overview of Data Science and Artificial Intelligence in the Field of Chemistry: "FS-08: Chemical Data Science and Artificial Intelligence Discussion, Work Shop" Kick-off Meeting Sponsorship

モデレーター: 湯田 浩太郎 Kohtaro Yuta (株式会社インシリコデータ In Slico Data, Ltd)

化合物分野で最近頻繁に聞かれる言葉として、**薬物デザイン、仮想スクリーニング、ドラグリポジショニング、毒性予測/評価、機能性化合物デザイン**等の言葉(手法)がある。更には**動物実験代替法研究分野**での展開も急速に進行している。これらの手法はインシリコ(コンピューター)上で展開される研究である。これら手法の実施基本技術として**データサイエンスや人工知能**が展開されてきた。また、これらの基本技術はバイオや医療分野でも中核技術となる。関連で、DB、ネットワーク、クラウド等も関与してくる。

今後、5年、10年先まで同じ研究手法を継続できる保証はない。研究の大きな流れは、研究分野の 異なる研究者との**分野間連携・融合や学際研究**となっている。このような将来の研究発展や変化に寄 与できる可能性を有するデータサイエンスや人工知能を武器と出来れば心強い。さらには、自分の研 究に付加価値を付けることも可能となる。

- ◆参加していただきたい方々:データサイエンスや人工知能の素人の方、及び専門の方
- 1. WET 研究者であるが、インシリコ関連技術を自分の研究に適用したい、あるいは興味を持っている研究者。従って、インシリコ関連技術に関しては初心者となる。
- 2. データサイエンスや人工知能の専門研究者であるが、アプリに関する知識が無い。 即ち、化学分野の基礎知識が無く、化学分野への展開が出来ない意味での初心者。
- ◆今回のチュートリアル内容の基本方針:化学研究分野での適用技術中心の解説 本チュートリアルは、研究を実施する観点で全体を俯瞰できるようにすることである。即ち、自分の 研究にデータサイエンスや人工知能を適用する適用技術中心に説明する。

実際にデータサイエンスや人工知能を適用するのに必要な内容、即ち、①適用時の全体およびデータの流れ、②個々の手法の役割や適用限界、③化合物情報をコンピュータで扱うためのアナログ情報からデジタル情報への変換技術、④数値データの扱い方、⑤サンプル収集の留意点、⑥データ解析結果の評価法等の実施適用技術を中心に全体を説明する。

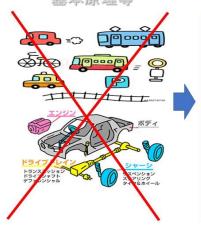
- ◆今回のチュートリアルの達成目標:
  - ・化学分野のデータサイエンスや人工知能の全体概要を知る。
  - ・WET 及び DRY 研究者間のコミュニケーションが取れて、共同研究を進められること。
  - ・将来の研究の変化に追随できるように化学データサイエンスおよび人工知能を知る。

### 化学関連研究の達成を、車で移動することに例えたイメージ図

今回はスキップ データ解析手法 基本原理等 今回のチュートリアル対象範囲

データ解析関連技術

適用分野







### 今回のチュートリアルでの討論内容概略



# チュートリアルの主要な内容

# 基本原理等

ニクラス分類 多クラス分類 重回帰

> ーラルネットワ**ナ**ク イナリーツリー

チャート分析

AdaBookst ランダムフォレスト その他 CNN RNN 深層学習

基本的な部分は概略説明

強化学習

## データ解析関連技術

#### ○データ解析の適用限界

- ・個々の手法に依存
- Oサンプリング
- 最小サンプル数
- ・クラスポピュレーション
- ・サンプリングプロトコル
- ・アウトライヤー、インライヤー
- ○線形/非線形問題
- ・空間合致と空間の再構築
- ・外挿と内挿
- Oパラメーター
- 種類
- ・パラメーター選択
- · 0値の扱い
- ・スケーリング
- ○解析信頼性
- ・サンプル数/パラメーター数
- ·要因解析
- ・クロスバリデーション
- ○その他

## 適用分野

#### ●創薬関連

- ·Q/T/ADME/P/SAR
- ・インシリコスクリーニング
- ・ドラグリストラクチャリング
- •要因解析
- ●化合物毒性評価
- ·TSAR(構造毒性相関)
- ·毒性予測
- ·脱毒性
- ·化合物&環境規制
- ●機能性化合物デザイン
- ·PSAR(構造物性相関)
- ●機器スペクトル
- ・スペクトル解析
- ・メタボロミクス
- ●バイオ解析関連
- ●医療との連携(画像等)
- ●その他