

MI²I チュートリアルセミナー (第9回)
「物質探索のための記述子設計」

「物質構造の表現と学習」

吉田 亮

情報・システム研究機構 統計数理研究所 データ科学研究系 教授
情報・システム研究機構 統計数理研究所 ものづくりデータ科学研究センター センター長
総合研究大学院大学 複合科学研究科 統計科学専攻 教授
物質・材料研究機構 特別研究員
〒109-8562 東京都立川市緑町 10-3
E-mail: yoshidar@ism.ac.jp



Liu Chang (劉 暢)

国立研究開発法人 物質・材料研究機構 NIMS ポスドク研究員
統計数理研究所 外来研究員
〒109-8562 東京都立川市緑町 10-3
E-mail: liu.chang@ism.ac.jp



MI の問題の多くは、順問題と逆問題の形式に帰着する。順問題の目的は、系の入力に対する出力の予測である。物性予測の文脈では、入力は物質（分子、組成、結晶など）、出力は物性値（エネルギー、電子状態など）に相当する。データ科学の文脈において、これらの解析は物質構造の“表現・学習・生成”を行うことに相当する。記述子と呼ばれる特徴ベクトルによって物質の構造を“表現”し、データのパターンから構造から物性の数学的写像を“学習”する。さらに、計算機を用いて所望の物性値を有する物質を“生成”し、有望な候補物質を炙り出す。対象となる入力は、分子、組成、結晶、混合物、プロセス、合成経路等、問題に応じて多様な形式をとりうる。

本チュートリアルでは、MI における最も基本的な要素技術である記述子に焦点を絞り、分子・組成・結晶構造を対象とする基本的な記述子に関する包括的レビューを行う。さらに「転移学習に基づく記述子の自動設計技術」や「記述子の設計とモデルの解釈可能性」等について解説する。また、Python に基づき開発された記述子計算用のソフトウェア「XenonPy」の紹介を行う。