

MI²I チュートリアルセミナー (第 10 回)
「マテリアルズ・インフォマティクスの基本原則と手順」

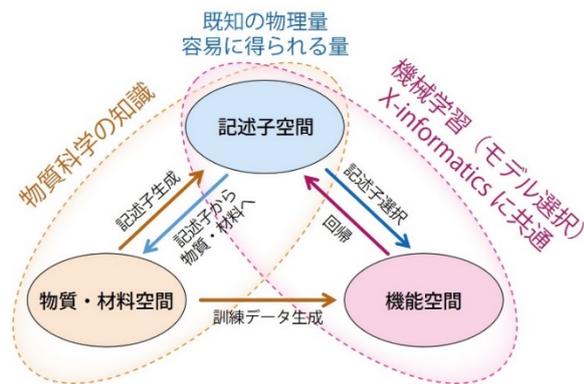
MI の基本原則は何か

寺倉清之 (産業技術総合研究所)

データ科学では、ある現象の観測データと、その現象に関係すると思われる特徴 (記述子) との間の相関関係を求め、それから未観測の現象の結果を機械学習によって予測する。図は、その様子を示す。従って、問題に応じた適切な記述子の選択と、機械学習手法の選択が重要になる。言うまでもなく、データ科学ではデータが基盤である。また、データ科学の中でもマテリアルズ・インフォマティクス (MI) においては、シミュレーションの果たす役割が非常に大きい。このような観点から、本稿では



- 1) 記述子に関する基本的な問題と
現状での関連活動状況
- 2) 機械学習の基盤となる Bayes 推定の
考え方と具体例および Bayes 最適化
- 3) データに係る問題と対応
- 4) シミュレーションの役割
について述べる。



職歴 :

1971 年 : 理学博士 (大阪大学) 、1978 年 : 東大物性研助教授、1990 年 : 同教授、その後工業技術院産業技術融合領域研究所、産総研、北大、北陸先端大を経て、2015 年から物質・材料研究機構フェローとして、情報統合型物質・材料開発イニシアティブ(MI²I)のプロジェクトリーダーを 2015 年から二年間、その後 2018 年から同プロジェクトのエグゼクティブアドバイザーを務めた。現在、AIST 名誉リサーチャー・エグゼクティブアドバイザー。専門は物性理論、計算科学、現在は物質科学への機械学習の活用に興味を持つ。

MI の手順 反復 5 ステップ

ダム ヒョウ チ （北陸先端科学技術大学院大学）

データ科学を材料科学研究に適用するには、以下の基本「5 ステップ」があることをしっかりと認識することが出発点と考えています。さらに、この 5 ステップは機械的に一巡すれば完了するものではなく、時として戻り、また繰り返しつつ進める必要があることも認識しておく必要があります。



- (1) 問題提起 = マテリアルサイエンスの問題をどのようにデータ科学の問題に焼き直すか？
(データ出発で求める結論を導くためのデータ科学的なシナリオ策定はこのステップに含まれます)
- (2) 提起した問題をデータ科学の手法で解決するためにデータをどのように適切に準備するか？
(記述子を適切に選択・設計するタスクはこのステップに含まれます)
- (3) データ科学の手法（アルゴリズム等）をどのように適切に適用するか？
- (4) 得られた結果の妥当性・有意義性をどのように評価するか？
- (5) 得られた結果をどのように材料科学分野のコミュニティに説明し、その展望について説得するか？

これらの 5 ステップを著者らが最近行った研究（希土類遷移金属二元合金の系キュリー温度の予測理解、結晶構造と特性相関理解のための記述子 OFM 開発、材料の類似性解析など）を事例として紹介する予定です。

参考論文

- Learning structure-property relationship in crystalline materials: A study of lanthanide transition metal alloys, *The Journal of Chemical Physics*, 148, 204106 (2018).
- Important descriptors and descriptor groups of Curie temperatures of rare-earth-transition-metal binary alloys, *Journal of the Physical Society of Japan*, 87, 113801 (2018).
- Committee machine that votes for similarity between materials, *IUCrJ*, 5, 830 (2018).
- Oxygen-diffusion-driven oxidation behavior and tracking areas visualized by X-ray spectro-ptychography with unsupervised learning, *Communications Chemistry Vol. 2*, 50 (2019).

職歴：

2003 年北陸先端科学技術大学院大学材料科学研究科博士課程修了。ハノイ大学 物理学科講師や北陸先端科学技術大学院大学知識科学研究科講師を経て、2011 年 4 月より北陸先端科学技術大学院大学知識科学研究科 准教授。専門はマテリアルインフォマティクス。主な研究テーマは磁石材料およびデータマイニング手法の開発。