

機械学習による高 T_c 超伝導物質探索 — 4 元系

Search for higher T_c superconductors by machine learning - quaternary system

中島弘幸、○松本 要、堀出朋哉 (九工大)

Hiroyuki Nakashima, ○Kname Matsumoto, Tomoya Horie (Kyushu Inst. Technol.)

E-mail: matsu@post.matsc.kyutech.ac.jp

I. Introduction

われわれは超伝導データベースと、人工知能技術の一種である機械学習で得られた予測モデルを活用することで、高 T_c を有する未知の超伝導物質を効率的に探索する新手法を開発している¹⁾。これまで3元系物質探索空間に限定し、NIMSのSuperConデータベース²⁾を基に、各種の材料データベースと機械学習技術を用いて3元系物質の T_c 予測モデルを構築してきた。その R^2 決定係数は0.92という高い精度を示し、このモデルを用いて予測した Mg-B-Ti 系や Fe-Te-Se 系の T_c 分布図は実験値とよい対応を示した。また、相平衡状態図との比較から合成が可能と思われる新しい高 T_c 物質候補を見出すことも可能となってきた。今回は、より T_c の高い銅酸化物超伝導体を含む4元系物質探索空間に範囲を広げて T_c 予測モデルの構築を試みたので、その結果について報告する。

II. Method

機械学習は統計解析手法を大量データに適用し、データ間に存在する有用な規則や分類を抽出する方法である。本研究では、既知物質と T_c の関係について NIMS の SuperCon データベースを利用した。今回は BaKBiO、YBCO、LaSrCuO 等を含む4元系超伝導物質として1200個程度を選択した。また材料データベースから、原子番号、原子量、電気陰性度、電子親和力、原子半径、イオン半径、イオン化ポテンシャル、電子状態、等々の53個の付随する記述子を収集してデータセットを作成した。 T_c の予測においては、データセットを訓練データと検証用のテストデータに80:20の比率でランダムに分割し機械学習を行った。予測データの偏りを抑えるために、このプロセスをランダムに10回繰り返しその平均値を最終値とする cross validation の手法を用いた。また、機械学習モデルとしては予備検討の段階で「回帰」「ラ

ンダムフォレスト」「サポートベクトルマシン」等の手法を試したが、本研究においては、もっとも予測精度の高い「ランダムフォレスト」による回帰手法を採用した。また学習モデルの予測精度の評価には R^2 決定係数を用いた。これは実測値と予測値の間の相関関係を表しており1に近いほど精度が高い。

III. Results and Discussion

Fig. 1 にランダムフォレスト回帰モデルによる予測 T_c と実測 T_c との相関関係を示す。10回の cross validation の結果、 R^2 決定係数は訓練データにおいて平均で0.89、テストデータでは0.87に達し、予測に使えるレベルの精度が得られていることが分かった。なお、データセットに含まれる最大の T_c は110 Kであった。今回のデータセットは銅酸化物の T_c データが主体となっているため、異方性の強い物質系の予測モデルになっていると考えられる。予測の詳細については当日報告する。

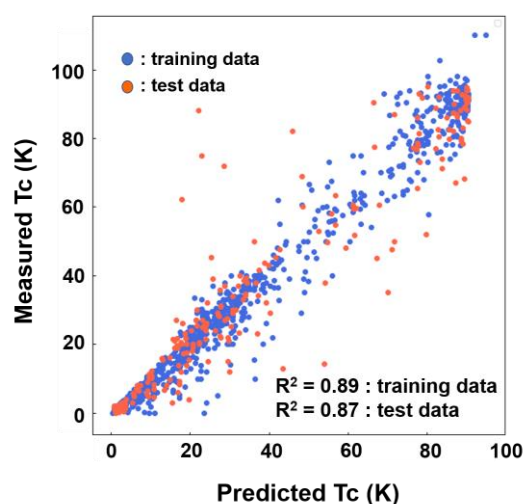


Fig. 1 4元系における予測 T_c と実測 T_c の相関

References

- 1) K. Matsumoto and T. Horide: Appl. Phys. Exp. **12**, 073003 (2019).
- 2) 超伝導データベース: <http://supercon.nims.go.jp/>