高 T_c 物質予測のための機械学習モデリング

Machine learning modeling for predition of higher T_c superconducting materials

°松本　要、堀出朋哉（九工大）
°Kaname Matsumoto, Tomoya Horie (Kyushu Inst. Technol.)
E-mail: matsu@post.matsc.kyutech.ac.jp

I. Introduction

一般に，高 T_c を有する新奇超伝導物質の探索は大変難しいが, かつては電子状態情報に基づくマティアスルールによってNb系等の超伝導物質の探索が進められた。銅酸化物や鉄系超伝導体の発見以降は，Cu-O 面や Fe-As 面の情報を手掛かりに，様々な結晶合成が行われ，多くの新奇超伝導物質が見出された。さらに，最近では水素化合物に超高圧を与えて常温に迫る T_c を持つ超伝導物質の探索も進められている。一方，近年の計算機パワーや計算アルゴリズムの進展に伴って,第一原理計算や従来の実験結果に基づいて,ハイスループットで新奇な超伝導物質を見出せる可能性が出てきている。本研究では, 機械学習によって従来探索されることのなかった 3 元系，4 元系，5 元系，等々の多元素系物質探索空間にその探索範囲を広げ，高 T_c を示す新しい超伝導物質を見いだすことを目的とする。われわれはこれまで，機械学習に基づく T_c 予測手法について報告してきた。ここでは，3 元系に焦点を当て，高い T_c を有すると期待される新奇超伝導物質のいくつかの例を紹介する。

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 の超伝導材料データベース[1] を利用した。今回は MgB_2 を含む 3 元系超伝導物質 1400 個を選択し，Materials Project のデータベース等[2] 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, 等々の約 150 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。この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 の比率で分割し機械学習を繰り返して行った。各学習モデルの評価には R^2 スコアを用いた。

III. Results and Discussion

当グループでは未知の多元系物質の T_c を予測する機械学習モデルをすでに作成済みで R^2 決定係数は 90-95%に達している。近年の水素化合物の高 T_c にみられるように，軽元素，例えば H，Li，Be，B，C 等を含む物質は高 T_c 新奇超伝導物質の候補である。図 1 に，B-H-Li と B-H-C 系の例を示す。われわれのモデルによれば，これらはどちらも 40 K に達する T_c を示す組成を持つ。課題は，その合成法や結晶

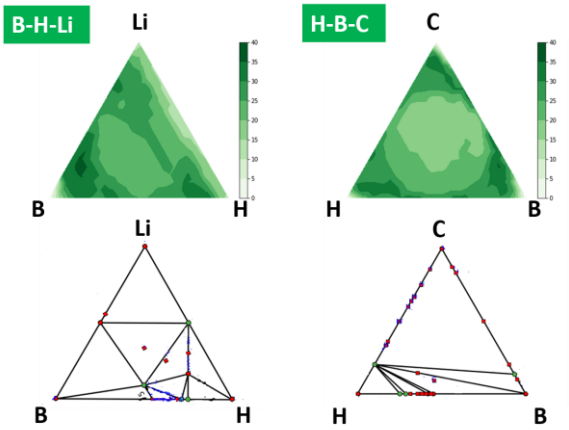

図 1　機械学習によって求めた H-B-Li 系および H-B-C 系における T_c 分布図と状態図

構造が不明なことである。今後，高圧合成や超高圧下の実験, および構造予測シミュレーション等でその特性を確かめることが期待される。

また，モデルの信ぴょう性については，訓練データから MgB_2 や Nb_3Sn 等の物質を除いても，機械学習モデルはこれらの物質の T_c を 90%以上の精度で予測できていることを付記しておく。これはもし MgB_2 発見以前であれば計算で MgB_2 を発見できた，ということを意味する。

References

1) 超伝導データベース: http://supercon.nims.go.jp/
2) Materials Project: https://materialsproject.org/