

教師なし学習で周期表を再発見できるか？ ～規則性の発見と視覚化～

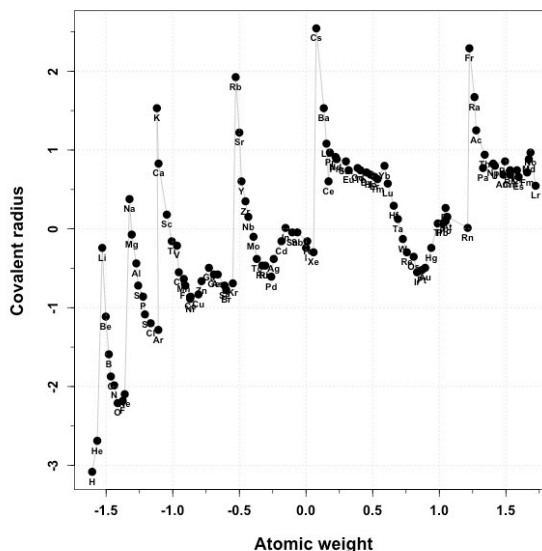
総合研究大学院大学	草場 穂
総合研究大学院大学	郭 中梁
統計数理研究所	Stephen Wu
統計数理研究所	吉田 亮

1. はじめに

周期表は、元素を周期律に従って配列した表である。現在の周期表の原型は、19世紀後半にロシアの化学者であるメンデレーエフによって作られた。当時、原子の構造はほとんど知られていなかったが、それぞれの元素が持ついくつかの特徴は知られていた。メンデレーエフはそれらの特徴が周期的に変化することを発見し、周期表を作り上げた[1]。この偉大な業績を機械学習の観点から見ると、多次元データの規則性を認知し、それを上手く視覚化するという、教師なし学習の問題だと考えることができる。そこで本研究では、当時知られていた元素の特徴データをもとに、周期表の再発見を試みる。

2. 研究意義

元素の特徴データを原子量順に並べると、明らかな規則性が確認できる(図)。それゆえ、一見簡単な問題に思えるが、意外に奥深いものがある。実際、主成分分析、カーネル主成分分析、t-SNE[2]など従来の次元圧縮法では上手くこの規則性を捉えることができなかった。また、先行研究でもあまり良い結果が得られていない[3]。その原因として、パターンの複雑さが考えられる。右の図をよく見てみると、複数のパターンの混合が読み取れる。これはヒトの眼には明らかだが、機械に認識させることは困難である。ゆえに、本研究の意義は、従来法では困難な規則性の認知と視覚化の問題を解決する手法を開発する事にあると言える。手法の詳細と結果は発表当日に報告する。



参考文献

- [1] 梶雅範 (1997) 『メンデレーエフの周期律発見』 北海道大学図書刊行会.
- [2] Maaten, L. and Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research* 9, 2579-2605.
- [3] Lemes, M.R. and Dal Pino, A. (2011). Periodic Table of the Elements in the Perspective of Artificial Neural Networks. *Journal of Chemical Education* 88(11), 1511-1514.