

MIにおける有機固体の記述法の検討

(早大データ科学¹・早大先進理工²) ○谷口 卓也¹・細川 真由子²・朝日 透²
 Examination of description methods of organic crystals in materials informatics (¹Center for Data Science, Waseda University, ²School of Advanced Science and Engineering, Waseda University) ○Takuya Taniguchi,¹ Mayuko Hosokawa,² Toru Asahi²

Organic crystals are composed of organic molecules, and molecular and crystal structures can be used in materials informatics (MI). There is a scarce report on the effectiveness comparison of molecular and crystal representations on the prediction task, and a fundamental question arises in MI of organic solids: which descriptor is useful and how effective it is in prediction task. This work aims to answer this question using currently available datasets where crystal structure and property are correlated. One dataset contains bandgap as reported in the literature, and the other is melting point, which we have curated from CSD database. Using these datasets, we compared regression results of molecular and crystal representations. Several fingerprints and graph were examined for molecular representation, and graph representation was examined for crystal structure. As the result, crystal graph afforded higher prediction accuracy for bandgap, while the regression results of melting point were almost same in all representations. These results will be discussed based on the difference of structure descriptions and targeted properties.

Keywords : Materials Informatics; Organic Crystal; Descriptor; Graph Network

有機固体は有機分子からなる結晶材料群であり、マテリアルズインフォマティクス (MI) において分子構造を記述子とすることも結晶構造を記述子とすることも可能である (下図)。しかし、有機固体のどの物性に対してどの記述法が適切かについては検討されていない。本研究では、分子構造の記述法としていくつかのフィンガープリントと分子グラフを用い、結晶構造の記述法としては結晶グラフを利用することで物性値に対する予測性能を比較・検討することとした。

目的変数がバンドギャップの場合、分子表現より結晶グラフを用いた方が予測性能は高く、目的変数が融点の場合、分子表現でも結晶グラフでも同等の予測性能となった。有機固体の記述法と対象とした物性の違いから得られた結果について議論する。

