Data-driven crystal structure prediction using structure similarity

([1]*SOKENDAI* [2]*The Institute of Statistical Mathematics*) ○Minoru Kusaba,[1] Chang Liu,[2] Ryo Yoshida[2]

Keywords: Crystal structure prediction; Metric learning; Crystal structure similarity; Atomic substitution; Computational materials discovery

The prediction of energetically stable crystal structures formed by any given chemical composition is one of the central problems yet to be addressed in solid-state physics. In principle, the crystalline state of assembled atoms can be found by solving the optimization of the energy surface, which can be evaluated by performing ab initio calculations. However, for a complex system, such as the one with a large number of atoms in the unit cell, the iterative gradient descent on the potential energy surface using first-principles calculations is prohibitively expensive. Here we present a unique methodology for the crystal structure prediction (CSP) that relies on a machine learning algorithm called metric learning[1)]. It is shown that the binary classifier, trained on a large number of crystal structures already identified, can determine the isomorphism of crystal structures formed by two given chemical compositions with an accuracy of about 96.4%. For any given query composition of unknown stability, the model can be utilized to automatically select from a crystal structure database a set of template crystals with nearly identical stable structures to which atomic substitution is to be applied. Apart from the local relaxation calculation of the predicted crystal structures, the proposed method does not use any ab initio calculations. The potential of this substation-based CSP is demonstrated on a wide variety of crystal systems. The numerical experiments with a large number of compounds (N=33,153) suggested that about 66.6% of their crystal structures are predictable by the CSP method.

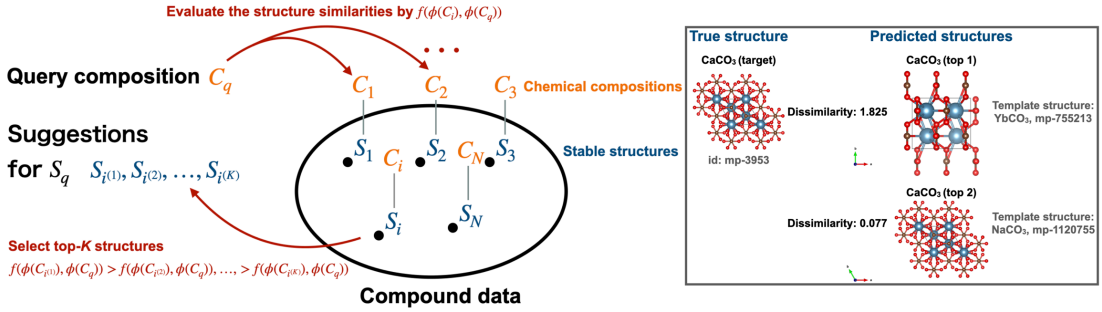

Figure 1: Schematic view of the substitution-based CSP based on metric learning (left), and Actual application example of the substitution-based CSP (Right).

1) B. Kulis, et al., Metric learning: A survey, Foundations and Trends in Machine Learning 5 (4) (**2013**) 287-364.