JAIST Repository
https://dspace.jaist.ac.jp/

Title	データマイニングを用いた量子計算データからの二元合金の物性予測
Author(s)	鈴木, 大輔
Citation	
Issue Date	2014-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/11994
Rights	
Description	Supervisor:Dam Hieu Chi, 知識科学研究科, 修士

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

Prediction of binary alloy property from the quantum calculated data by data mining

Daisuke Suzuki

School of Knowledge Science

Japan Advanced Institute of Science and Technology

March 2014

Keywords: data mining, quantum calculation, LASSO, graph.

In this paper, I propose a method that combines deduction (quantum calculation) and induction (data mining) for material design. The essential idea of the method is a process consisting of sparse regressions and cross-validation for analyzing data of the materials. Further, by performing LASSO in parallel, I build a directed graph in which nodes are features and edges are relations between features, thus representing global image of the relation between the features. To demonstrate the effectiveness of the proposed method, I worked on two issues, prediction of binary alloy's melting point and description of relationship between the features in the prediction model by graph.

In the prediction of binary alloy's melting point, I represented binary alloy data by experimental/basic physical property data and quantum calculated data of molecular model composed of two and three atoms, and

Copyright © 2014 by Daisuke Suzuki

created a binary alloy database from them. Then, I performed multiple regression analysis, LASSO, on the database, and got the sparse linear regression model which predict melting point. I applied proposed method to the four alloy groups, alkali metal group, alkali-earth metal group, transition-metal group, and rare-earth metal group, and I evaluated the results. As a result, I got exact prediction model which score is over 0.90 in alkali metal group and alkali-earth metal group. But *score* of transition-metal group and rare-earth metal group were 0.87, 0.71 respectively. This means that the data prepared for this experiment isn't enough to describe the behavior of transition-metal group and rare-earth metal group. So it is expected that adding features which describe the behavior of the d-electron system improve the prediction accuracy. On the other hand, it is noteworthy that I could predict the exact melting point from such a simple model's data.

Subsequently, I extended the melting point prediction model and built a directed graph from it. Specifically, I performed LASSO in parallel and summarized the results, then built a graph. In the graph, nodes are features and edges are the relationship between features. Further, I measured importance of features in the prediction model by prediction risk and *score*.

I have confirmed from the graph that electric charge transfer is important factor in terms of prediction of melting point in both alkali metal alloy group and alkali-earth metal alloy group. And I have confirmed the difference of nature with respect to component B too, which is while mass is important for alkali metal group, number of valence electrons is important for alkali-earth metal group. On the other hand, I understood that there weren't enough features to predict melting point in transition metal group and rare-earth metal group. In this way, I could obtain information or knowledge from the graph, and this information or knowledge are difficult to obtain by performing LASSO simply.

From the above results, it can be said that proposed method is effective for material design.