Computational Materials Science 138 (2017) 135–148

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

New methods for prediction of elastic constants based on density functional theory combined with machine learning

Juan Wang [a,b], Xiaoyu Yang [a,b,*], Zhi Zeng [c], Xiaoli Zhang [c], Xushan Zhao [a], Zongguo Wang [a]

[a] Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China
[b] University of Chinese Academy of Sciences, Beijing 100049, China
[c] Institute of Solid State Physics, Chinese Academy of Sciences, Hefei 230031, China

ARTICLE INFO

Article history:
Received 10 February 2017
Received in revised form 9 June 2017
Accepted 13 June 2017

Keywords:
Prediction of elastic constants
Materials informatics
DFT calculation
Neural network
General regression neural network
Support vector regression

ABSTRACT

Elastic constants play critical roles in researching mechanical properties, but they are usually difficult to be measured. While density functional theory (DFT) calculations provide a reliable method to meet this challenge, the results contain inherent errors caused by various approximations. The data-driven approach of machine learning also laid a foundation for predicting material properties. In order to increase the accuracy of theoretical calculations results, in this paper we investigate using machine learning methods to both correct the elastic constants by DFT calculation, and to directly predict elastic constants. The single-hidden layer feedforward neural network trained by back propagation algorithm (SLFN), general regression neural network (GRNN) and support vector machine for regression (SVR) techniques are employed to build regression models to correct the elastic constants by DFT calculation for metal or metallic binary alloys. We also build regression models to predict the elastic constants of metallic binary alloys with cubic crystal system rather than using DFT calculations. It has been demonstrated that the elastic constants corrected by regression models has higher accuracy than those calculated by DFT, and the elastic constants of binary alloys directly predicted by model using the outperformed SLFN technique is prospective.

1. Introduction

The density-functional theory (DFT) [1] provides an efficient approach for materials design and microscopic mechanism analysis from computing perspective. However, due to the inherent approximations adopted in DFT, there is a deviation between a DFT prediction and the corresponding experimental value. Despite the DFT prediction suffer from lack of precision, the calculation results also could capture the essence of the physics, such as electronic structure, behavior of lattice vibrations, optical property, magnetic moment, anisotropy elastic constants, and so on. In order to reduce the deviation and increase the accuracy of DFT calculations, in this study, an approach of using DFT calculation together machine learning is used to minimize the gap between theoretical values and the experimental measurements. Apart from improving DFT calculation results, the machine learning is also used to predict anisotropy elastic constant of a binary alloy by using relevant data of its comprising pure elements.

Material properties greatly depend on chemical composition and crystal structures. For the same property, materials with different elements, composition and structure often show different values. It means that the property value of a material has relationships with its comprising elements, composition and structure. Regression techniques have laid a foundation for modeling structure-property relationships in material informatics [2–4]. These methods take a set of known material structure information along with some properties (e.g. strength, band gap, melting temperature, *etc.*) as inputs, and output the predicted value. In the past decade, machine learning methods have been successfully applied in the field of material informatics to predict material properties and to improve the accuracy of DFT calculations. For example, in 2004, Guanhua Chen applied neural networks to improve the accuracy of DFT calculation of heat of formation of 180 organic molecules, and the error of the corrected heats of formation were dramatically reduced [5]. In 2009, Ozerdem and Kolukisa applied artificial neural network to predict the mechanical properties of Cu–Sn–Pb–Zn–Ni cast alloys using chemical compositions [6]. In 2016, Balachandran used Gaussian Process Model (GPM) and SVR to train a model that predicts elastic properties in terms of elementary orbital radii of the individual components of compound mate-

* Corresponding author at: Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China.
E-mail address: kxy@cnic.cn (X. Yang).

http://dx.doi.org/10.1016/j.commatsci.2017.06.015
0927-0256/

rials [7], and the initial training data they used are calculated by DFT [8]. In 2016, Ting Gao used general regression neural network (GRNN) to correct the DFT non-covalent interactions calculations based on the benchmark databases S22, S66 and X40 [9]. It has shown that the data-driven machine learning methods can well capture the structure-composition-property relationship, and used to reduce calculation errors caused by inherent approximations in the level of theory and limited basis sets.

In this study, we investigate using machine learning to both correct the elastic constants value by DFT, and directly predict the elastic constants for binary alloys. In order to get prediction value of elastic constants with higher accuracy for metal elements and binary alloys, we use machine learning to develop a model by using dataset of DFT calculation of elastic constants, dataset coming from crystal structure information, and composition information to improve the accuracy of DFT calculation elastic constants. That is to say, we take composition information, DFT calculations of volume, energy, and elastic constants as inputs for regression model, and the model outputs the corrected elastic constants of DFT calculation with different DFT XC-functionals. In terms of direct prediction of elastic constants for binary alloys, we build a machine learning model using crystal structure information, composition information, and experimental elastic constants of its comprising pure elements to predict the elastic constants of binary alloys. It has been discovered that the model prediction values are closer to the experimental values compared with DFT calculation values.

The procedures of the proposed approach are described as follows: (1) We collect experimental data of elastic constants for 34 metal pure elements and 59 binary alloys from literatures, which are used to tuning parameters of regression models, and use DFT software package to produce simple physical properties (e.g. volume, energy) and elastic constants of these 93 materials with different input settings of DFT calculation; (2) We briefly introduce three kinds of machine learning techniques, SLFN, GRNN and SVR to build regression models, and describe the predictors of the regression models; (3) Based on collected data set, the SLFN, GRNN and SVR are used to construct models to correct DFT calculation elastic constants of pure element metals and binary alloys. The results by the three regression models are compared. (4) At last, as a test case of predicting property straightaway rather than using DFT, the elastic constants of binary alloys are predicted using experimental elastic constants and crystal structure information of its comprising pure elements.

The remainder of this paper is organized as follows. Section 2 describes the data preparation for using machine learning technique, and discusses three machine learning methods as well as predictors for building regression model. The Section 3 focuses on the analysis of regression and application performance of the proposed approach. Section 4 is the conclusion.

2. Methods

2.1. Data preparation

The prediction models are built from a data set containing experimental elastic constants and predictors for 34 pure element metals and 59 binary alloys. The predictor values of physical properties, such as volume, energy and elastic constants are obtained from DFT calculation. In the present work, DFT calculations are performed by the VASP code [10,11]. The ion-electron interaction is described by the projector-augmented wave (PAW) method. The used XC functionals are depicted by the generalized gradient approximation (GGA) parameterized by Perdew and Wang (PW91), Perdew Burke Ernzerh (PBE), and Local Density Approximation (LDA). The applied plane-wave kinetic energy cut-off value is the largest found among the recommendations for all species involved in the calculation, increased by a factor of 1.3. The convergence test of total number of k-points (KPT) in the irreducible Brillouin zone for all DFT calculation has been done in this paper, and when the KPT is set to 21,000, the DFT calculation results for the sample materials can reach a stable convergence state. So in our work, the number of KPT is set to 21,000. The k-point mesh can be generated using Monkhorst-Pack scheme according to the value of KPT as well as the reciprocal lattice vector of the simulation cell and the total number of the atoms in the simulation cell. The DFT calculation and experimental elastic constants of these 93 materials are list in the Appendix.

To get the predictor values, three kinds of DFT calculation types: RELAX calculation, STATIC calculation, and ELASTIC CONSTNTS calculation should be carried out. There are 3 sets of DFT calculation parameters, and it needs to run 837 times of VASP code to get physical properties value of predictors and elastic constants of these 93 materials. The data extraction and data management for all the calculations is also tedious. By using the high-throughput DFT simulation engine and data management platform, namely MatCloud [12], this research study is doable and workload is greatly reduced. Among the elastic constants calculated by the three types of pseudopotential PAW_PW91, PAW_PBE, PAW_LDA for each material, we use the elastic constants with the minimum error as predictor, and the values of other simple physical properties predictors are obtained from the DFT calculations with the same pseudopotential as that used for ELASTIC CONSTANTS calculation.

2.2. Machine learning methods

Regression is a useful machine learning technique [13] to construct a model that predicts response variables from a set of independent variables. Least Squares Regression (LSR) is a commonly used method to estimate regression models. It can effectively describe linear relationship between response variables and independent variables. However, the prediction from LSR is not very good in nonlinear regression applications [14]. The general regression neural network (GRNN), SLBPFN and SVR are nonlinear regression methods that show good prediction ability in physical and material study fields [15,16]. These three learning algorithm are stated as follows.

2.2.1. Single-hidden layer feedforward neural network (SLFN)

SLFN has inherent learning and generalization abilities. It has been used in a variety of applications successfully. SLFN learns from examples, much like human beings. A SLFN with a non-polynomial activation function can approximate any continuous function to any degree accuracy [17]. A standard SLFN with at most N hidden neurons and with any bounded nonlinear activation function which has a limit at one infinity can learn N distinct samples with zero error [18].

Consider a data set of N different data $D = (\boldsymbol{x}_i, \boldsymbol{t}_i)$, where $\boldsymbol{x}_i = [x_{i1}, x_{i2}, \ldots, x_{io}]^T \in R^o$, $\boldsymbol{t}_i = [t_{i1}, t_{i2}, \ldots, t_{im}] \in R^m$, $\boldsymbol{x}_i$ and $\boldsymbol{t}_i$ denote a vector of O predictor variables and M response variables. The mathematical model of SLFN with $\bar{N}$ hidden nets can be described as follow

$$\sum_{i=1}^{\bar{N}} f^2(\boldsymbol{w}_i^2 \boldsymbol{x}_{ij}^2 + \boldsymbol{b}^2) = \sum_{i=1}^{\bar{N}} \sum_{k=1}^{O} f^2(\boldsymbol{w}_i^2 f_i^1(w_{ki}^1 x_{kj} + b_i^1) + \boldsymbol{b}^2) = \boldsymbol{t}_j, \quad j = 1, 2, \ldots, N \tag{1}$$

where vector $\boldsymbol{w}_i^2 = [w_{i1}^2, w_{i2}^2, \ldots, w_{im}^2]^T$ denotes the weights that connect the i-th hidden net and the output nets, function $f^2(x)$ is the