Proposal - Transparent Conductor Prediction

Yan Sun yas108@eng.ucsd.edu
Yiyuan Xing yix016@eng.ucsd.edu
Xufan Xiong x7xiong@eng.ucsd.edu
Tiaoduo Hao t6hao@rng.ucsd.edu

1 Introduction

Nowadays the development of innovative materials is one of then most challenges for physical application, which concerns about the development of heath equipment, new energy application and many other fields. In order to optimize the property of materials, it is crucial to get a deep understanding the relationship among properties, composition and internal energy condition. Specifically, transparent conductors are significant compounds that are electrically conductive and low absorption in the visible range, which is a special property of these conductors and could make them applied to sensors, transistors and laser equipment[2].

However, one of the biggest problems is only a small portion of compounds is well understood that is able to be considered as transparent conductor. In order to find the optimum composition for transparent conductor, some basic principle should be the basic rule for computational approach. The alloys for transparent conductor should be $(Al_xGa_yIn_z)_{2N}\ O_{3N}$, where $x + y + z = 1$ and N is an integer (usually between 5 and 100). There are infinite possible combinations for the values of x,y and z so the choice of computational method is the pivotal issue for transparent conductor materials design efficiency. There exist one primary computational method for materials science named Density Functional Theory, which is able to get high accuracy result but requires much computing time even for supercomputers. In this way, the data-driven method will be an alternative way to improve the efficiency for the transparent conductor design process.

2 Dataset and Method

There exist usable data for transparent conductor online[1]. In order to assist in the transparent conductor, the specific task in this assignment is to predict two significant properties, which is formation energy and bandgap energy. Formation energy is related to the stability of materials and bandgap energy is related to the transparency over the visible range. This prediction procedure will be considerably important for aforementioned design process.

Each team member will try to focus on one type of method to make the prediction then all the models will be compared. Initially, Yiyuan decides to try Random Forest method, Xufan trys Neural Network method, Yan make experiment on Adaboost and otherensemble learning method and Tianduo focuses on Catboost method. The specific plan may be changed based on practical progress. The objective of this project is to build models that make RMSLE result of prediction result as low as possible.

References

[1] Kaggle: predicting transparent conductors. https://www.kaggle.com/c/nomad2018-predict-transparent-conductors/data.

[2] D. S. Ginley and J. D. Perkins. Transparent conductors. *Handbook of Transparent Conductors*, page 1–25, Nov 2010.