Summit to: Computational Material Science

# How to determine universal formula for given database

**Bo Da*, Jiangwei Liu, Hideki Yoshikawa, Shigeo Tanuma**

National Institute for Materials Science, 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan

**EXTENDED ABSTRACT:** The robust TPP-2M formula is the most popular empirical formula for the prediction of charged particle inelastic mean free paths from simple material parameters. However, the TPP-2M formula poorly describes several materials because it adopts traditional least-squares analysis. Herein, we propose a new framework based on machine learning. This framework allows a selection from an enormous number of combined terms (descriptors) to build a new formula. The number of terms in the new formula can be automatically adjusted according to the importance of the terms in the particular application scenario. The obtained framework not only provides higher average accuracy and stability but also reveals the physical meanings of several newly found descriptors, and by the principle descriptors found, a complete physical picture of IMFP is summarized. Our findings suggest that machine learning is powerful and efficient and has great potential in building a regression framework for data-driven problems as shown in Fig. 1.

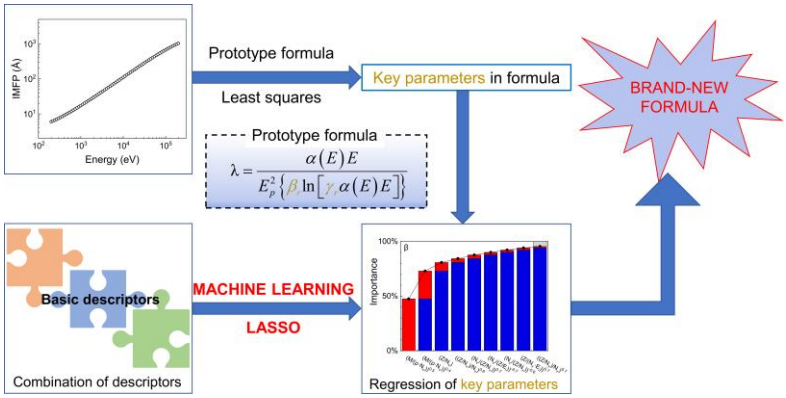


**Figure 1**. The flowchart of the framework.

Besides the robust TPP-2M formula is the most popular empirical formula for the prediction of charged particle inelastic mean free path, there is also some popular empirical formula for the prediction of secondary electron yield in surface analysis. Knowledge of absolute secondary electron yield ($\delta$) is important for various applications of electron emission materials. Besides, it is also crucial to know the dependence of $\delta$ on primary electron energy $E$p and material properties like atomic number Z. The available experimental database of $\delta$ reveals a large discrepancy among the measurement data, while the oversimplified semi-empirical theories of secondary electron emission can only present the general shape of the yield curve but not the absolute yield value. This limits not only the validation of a Monte Carlo model for theoretical simulations but also presents large uncertainties in the applications of different materials for various purposes. In applications, it is highly desirable to have the knowledge of the absolute yield of a material. Therefore, it is highly desirable to establish the relationship of the absolute yield with material and electron energy based on the available experimental data. Recently, machine learning (ML) methods have been increasingly used for the prediction of material properties mainly based on the atomistic calculations with the first-principles theory. We propose here the application of ML models to a material property study, starting with experimental observations

* Corresponding author: DA.Bo@nims.go.jp.

and unfolding the relationship of $\delta$ with basic material properties and primary electron energy as shown in Fig. 2. Our ML models can predict ($E$p)-curve covering a wide energy range of 10 eV–30 keV for unknown elements within the uncertainty range of the experimental data and can suggest more reliable data among the scattered experimental data.

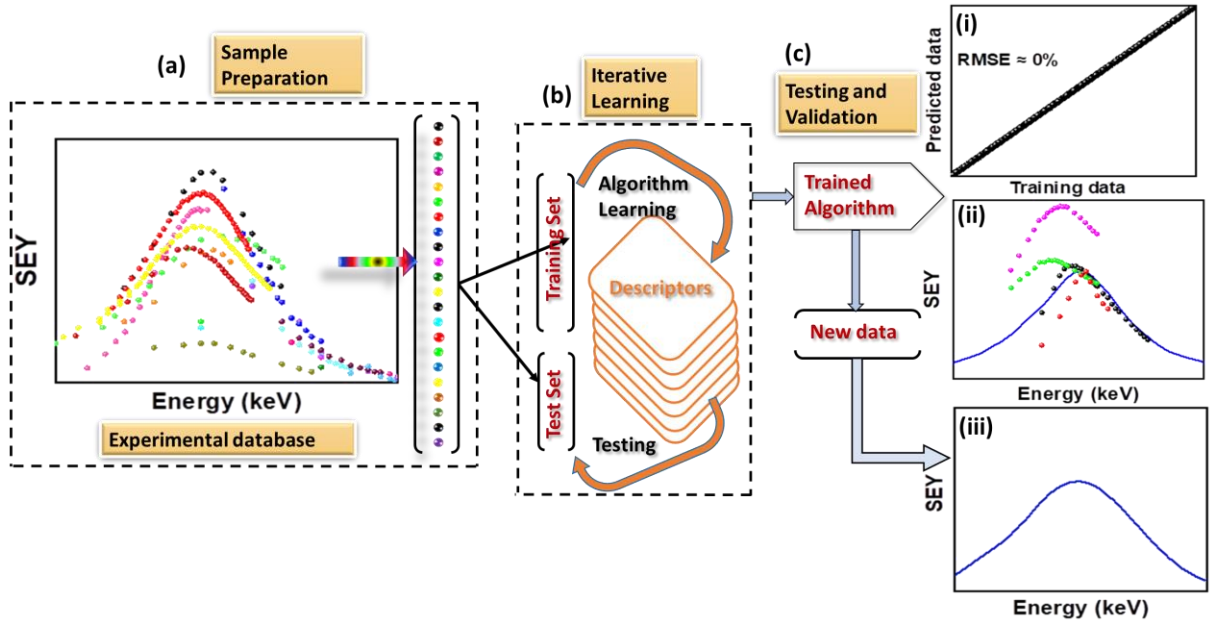


**Figure 2**.Using Machine Learning methods on the noisy experimental data to predict reliable SEY for elements in the periodic table.

In addition, we extended the model to predict the yield for 20 elements (in total) beyond the training dataset, including 6 elements (Ni, Cd, Hg, Tl, La, Th) for which experimental data exist and 14 elements (Na, Mn, Co, As, Rb, Ru, Rh, Os, Ir, Ce, Nd, Sm, Eu, Lu) for which no a measurement was performed. We conclude here that ML is an effective tool for the prediction of secondary electron yield. In the future, this approach can also be extended to explore other material properties from available experimental database. Furthermore, the obtained SEY data will be helpful to derive a mathematical expression for the calculation of SEY, which is very much needed by the scientific community.

**Keywords:** charged practice solid interaction, machine learning, inelastic mean free path, database;

**REFERENCES**

[1] Xun Liu; et al. Physical Chemistry Chemical Physics. **25** (2023) 17923.

[2] Mehnaz; et al. Physical Chemistry Chemical Physics. **25** (2023) 17238.

**BIOGRAPHY**


In 2008 Dr Bo Da obtained a BS in Physics from University of Science and Technology of China (USTC) and in 2013 a PhD in physics from the same university. In November 2013 he moved to the National Institute for Materials Science (NIMS) (Tsukuba, Japan) as a Postdoctoral Research Fellow, in January 2015 becoming an ICYS Researcher at their International Center Young Scientists (ICYS), in December 2015 becoming a Researcher in the Center for Materials Research by Information Integration (Mi2) and promoted as Senior Researcher in April 2019.