DOMAIN KNOWLEDGE-GUIDED MACHINE LEARNING AND CASE STUDIES OF METAL OXIDATION

Tong-Yi Zhang[1,2]*

[1] Guangzhou Municipal Key Laboratiry of Materials Informatics, Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China [2] The Materials Genomy Institute, Shanghai University, Shanghai, Chian

* *Presenting Author Email: zhangty@shu.edu.cn*

Abstract

This presentation first briefly introduces the concept of materials/mechanics informatics, which integrates machine learning with materials/mechanics science and engineering to accelerate materials/mechanics, products and manufacturing innovations. Then, this presentation reports a domain knowledge-guided machine learning strategy and demonstrates it by studying the oxidation behaviours of ferritic-martensitic steels in supercritical water and oxidation behaviours of FeCrAlCoNi based high entropy alloys (HEAs) at high temperatures. This strategy leads to the development of formulas with high generalization and accurate prediction power, which are most desirable to science, technology, and engineering.

1. Introduction

Theories and models in fracture and failiure of materials and structures are usually represented by mathematical formulas, which play vital roles in the understanding of fracture and failiure behaviors. Before machine learning is available, theories and models are developed by human beings through intuition, derivation, and/or summarization from the accumulation of experience and knowledge of fracture and failiure of materials and structures. This model-developing paradigm, however, requires long time and huge efforts, in particular the geniuses of human beings. Nowadays, the rapid development of artificial intelligence (AI) techniques provide an opportunity for automatic model construction. The new paradigm of AI for materials/mechanics by using AI techniques such as machine learning (ML) to discover knowledge and science from data generated from both experiments and computational simulations prevails rapidly in the field of materials/mechanics science and engineering. The AI for materials/mechanics paradigm, combined with expert domain knowledge, provides state-of-the-art methodologies for model development. Without using domain knowledge, the search space for analytic formulas is usually too huge to find scientifically meaningful equations. Therefore, the stratagy of domain knowledge-guided machine learning is proposed to accelerate the data-driven discovery of theories and models of fracture and failiure of materials and structures. The oxidation data of ferritic-martensitic steels in supercritical water and the oxidation data of FeCrAlCoNi based HEAs at high temperatures were collected from the literature. These experiments were condicted by various research groups so that the data are diversified and it is a challenge to find a general analytic formula for each of the two systems. Domain knowledge-guided machine learning strategy is able to tackle the challenging problem, where domain knowledge suggests a time-dependent (power law) Arrhenius equation. Accordingly, the Tree-Classifier for Linear Regression (TCLR) algorithm is developed to solve the challenging problem.

2. Results

As an example, the dimensionless Arrhenius equation for the oxidation of FeCrAlCoNi based HEAs in air at high temperatures is given here: $\frac{\Delta w}{\Delta w_{00}} = \frac{\Delta w_0}{\Delta w_{00}} \left(\frac{t}{t_0}\right)^m \exp\left(-\frac{Q}{RT}\right)$, where $\Delta w(\mathrm{mg/dm^2})$: weight gain, $t(h)$: time, m: time exponent, T(K): absolute temperature, R(8.314 J/mol·K): gas constant, Q(J/mol): activation energy, and $\Delta w_{00} = 1\mathrm{mg/m^2}$ and $t_0 = 1\mathrm{h}$ denote the reference weight gain and time, respectively; $\frac{\Delta w_0}{\Delta w_{00}}$ is named the dimensionless pre-factor, and $\frac{t}{t_0}$ is the dimensionless time. The dimensionless Arrhenius equation for the oxidation of ferritic-martensitic steels in supercritical water is basically the same. Figure 1 shows how the strategy works.

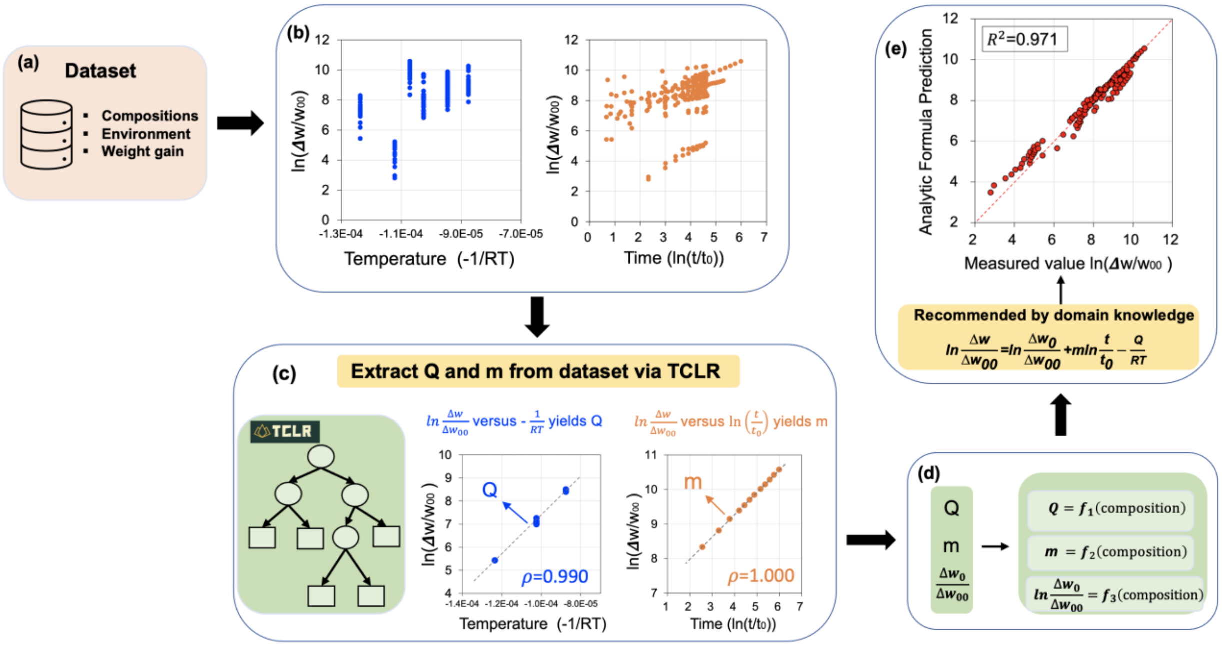

Fig.1. Schematic diagram of the workflow. (a) The collection of high temperature oxidation data of FeCrAlCoNi based HEAs in air from the literature, (b) the diversity of the original data, (c) TCLR giving high linearity between $\ln\left(\frac{\Delta w}{\Delta w_{00}}\right)$ and $-\frac{1}{RT}$ and between $\ln\left(\frac{\Delta w}{\Delta w_{00}}\right)$ and $\ln\left(\frac{t}{t_0}\right)$ for the data in each of leaves, (d) the three spectrums as functions of material composition, and (e) the prediction of the final analytic formula.

The explicit dimensionless Arrhenius equation includes

$$Q^{\mathrm{LR}} = 123.015\mathrm{Al}^2 + 66.051\mathrm{Cr}^2 + 66.051\mathrm{Fe}^2 + 178.086\mathrm{Co}^2 + 66.051\mathrm{Ni}^2 + 8198.492\mathrm{Cu} - 229.703\mathrm{Cu}^2 - 20753.850\mathrm{Ti} + 967.148\mathrm{Ti}^2 + 929.394\mathrm{Si}^2 (\mathrm{kJ/mol}),$$

$$m^{\mathrm{LR}} = 579.51\mathrm{Al} - 0.069\mathrm{Al}^2 + 608.80\mathrm{Cr} - 0.964\mathrm{Cr}^2 + 608.80\mathrm{Fe} - 0.964\mathrm{Fe}^2 + 542.86\mathrm{Co} + 1.455\mathrm{Co}^2 + 608.80\mathrm{Ni} - 0.964\mathrm{Ni}^2 + 579.50\mathrm{Cu} - 0.068\mathrm{Cu}^2 + 593.39\mathrm{Ti} - 1.596\mathrm{Ti}^2 + 94.316\mathrm{Si}^2 - 58372.892,$$

$$\ln\left(\frac{\Delta w_0}{\Delta w_{00}}\right)^{LR} = -1353.273\mathrm{Al} + 0.179\mathrm{Al}^2 - 4268.779\mathrm{Cr} + 2.320\mathrm{Cr}^2 + 2.320\mathrm{Fe}^2 - 1262.339\mathrm{Co} - 3.604\mathrm{Co}^2 + 2.320\mathrm{Ni}^2 - 1352.441\mathrm{Cu} + 0.143\mathrm{Cu}^2 - 1389.475\mathrm{Ti} + 4.030\mathrm{Ti}^2 - 220.136\mathrm{Si}^2 + 136299.812.$$

3. Conclusion

The present work demonstrates the significant role of domain knowledge in the development of materials/mechanics informatics.

Acknowledgements

Mr. Bin Cao, Mr. Qinghua Wei, Mr. Lucheng Deng, Mr. Shuang Yang, Mr. Ankang Sun, Dr. Ziqiang Dong participated in this work. TYZ thanks Guangzhou government for the support to the Guangzhou Municipal Key Laboratiry of Materials Informatics.